

QY 180 CACAATCAATTAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTTACCCA 239
Dd 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGAAGCGGGAGACATGAGGTGGAATACTCCTGGAAGGAGGCCGTGTGCAGCG 299
Dd 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80
QY 300 GTCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAATATTAACATTTGCGGTGAACCTG 359
Dd 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Dd 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGCTGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG 479
Dd 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGAAAATGGCACCGGCCAAAGCCATCATACGCTCTCCCTGATGGGAAA 539
Dd 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACACCCCGGATGGAGAGATGGAATTTTCATCTACGCTCTCCACACACTT 599
Dd 161 PropheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAAGTATTTCCAGAAAATTTGGACGATGTTTCACTGAGAGTTTCTGTGAACACAGCCAAT 659
Dd 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCTCAACTCATGGAAGTACTGTCTACAGAGACATGACGGGCATAT 719
Dd 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAGTGAAGATGTGTACGTGTTAAACAGATCAGATTCCTGTGTTGTG 779
Dd 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTGGC-AAAGATCTCCCC 838
Dd 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTGTATGCTCTGATTCATGATCCTAGCCACTTCTCAATTTATTTCTACCATTAAC 898
Dd 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGATAAATACTGGCTGTTTGTGTTTCCACCAATCATCTGTGAAT 958
Dd 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGACCA 1018
Dd 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCGCCACCGCCACCCACCCAGACCTTCAAAACCCACCCCTTCTTTAGGA 1078
Dd 321 GlyProCysProProProProProProProProProProProProProProProProPro 340
QY 1079 CTTGCTGGTGACAAACCCCTGAGCTGAGTAGGATTCCTGATGAAAACTGCCAGATTAAAC 1138
Dd 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTCAAGCCACCATCAAAATTTGTAGAGGAATCTTAGAGGTTAACATC 1198
Dd 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGTCTGATGCGGTGCCATGGCCTGAAGCTCCCTAATAGACTTT 1258
Dd 381 IleGlnMetThrAspValLeuMetProValProTyrProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACCATCATTTCTGACCCACC 1318

Dd 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleSerAspProThr 420
QY 1319 TCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTG 1378
Dd 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACCTTCAATGGTCTGGACGTACTGTGTGAACCTCACCTGGGGAT 1438
Dd 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACAAAGCCTGGCTCTCACGACGACCTGATTTCTGTCTCTGACAGACCCAGCCTCG 1498
Dd 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGCTGGCCATATTTGTCACT 1558
Dd 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCTCTTGGTGTACAAAAACACAAAGGAATACAAACCAATAGAAAATAGTCT 1618
Dd 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGAATGTGTCAGAAAGCAAGGCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTT 1678
Dd 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGGAACACGAGAAAGGATCCGCTACTCAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Dd 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 2
S53871
Pmel 17 protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 05-Nov-1999
C;Accession: S53871
R;Kwon, B.S.; Halaban, R.; Ponnazhagan, S.; Kim, K.; Chintamaneni, C.; Bennett, D.; Pic
Nucleic Acids Res. 23, 154-158, 1995
A;Title: Mouse silver mutation is caused by a single base insertion in the putative cy
A;Reference number: S53871; MUID:95175358; PMID:7870580
A;Accession: S53871
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-626 <KWO>
A;Cross-references: GB:U14133; NID:g887940; PIDN:AAA69538.1; PID:g887941

Alignment Scores:
Pred. No.: 1-49e-46 Length: 626
Score: 635.00 Matches: 184
Percent Similarity: 41.45% Conservative: 97
Best Local Similarity: 27.14% Mismatches: 207
Query Match: 13.09% Indels: 191
DB: 2 Gaps: 22

US-10-039-272-1 (1-2661) x S53871 (1-626)
QY 78 TTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTGATGCC-----GCCAAACGA 131
Dd 9 PheLeuProValLeuValLeuSerAlaLeuLeuAlaValGlyAlaLeuGluGlySerArg 28
QY 132 TTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTA 191
Dd 29 AsnGlnAspTrpLeuGlyValProArgGlnLeuVal----- 40
QY 192 AATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGG 251
Dd 41 -----ThrLysThrTrpAsnArgGlnLeuTyrProGluTrpThr--- 53
QY 252 GGACATGAGGTGGAAAACTCTTGAAGGGAGGCCCGGTGTGCAGGCGGTCTGACCACT 311
Dd 54 ---GluValGlnGlySerAsnCysTrpArgGlyGlyGlnValSerLeuArgValIleAsn 72

Next Scan pag 7, 8, 9

Db 1260 TGCCATGGCCTGAAAGCTCCCTAATAGACTTTTGTGCTGACCTGCCAAGGGAGCAATCCCA 1319
QY 1287 CGGAGGTCTGTACCATCATTTCTGACCCCACTCGGAGATCACCCAGAACACAGTCTGCA 1346
Db 1320 CGGAGGTCTGTACCATCATTTCTGACCCCACTCGGAGATCACCCAGAACACAGTCTGCA 1379
QY 1347 GCCCTGTGGATGGATGAGATGTGTCTGCTGCTGACTGTGAGACGAACCTTCAATGGGTCTG 1406
Db 1380 GCCCTGTGGATGGATGAGATGTGTCTGCTGCTGACTGTGAGACGAACCTTCAATGGGTCTG 1439
QY 1407 GGACGTACTGTGTGAACCTCACCCCTGGGGATGACACAAGCCTGGCTCTCACGAGCACCC 1466
Db 1440 GGACGTACTGTGTGAACCTCACCCCTGGGGATGACACAAGCCTGGCTCTCACGAGCACCC 1499
QY 1467 TGATTTCTGTTCTTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCCTTGA 1526
Db 1500 TGATTTCTGTTCTTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCCTTGA 1559
QY 1527 TCTCCGTTGGCTGCTTGGCCATATTTGFCACTGTGATCTCCCTCTGGTGTACAAAAAC 1586
Db 1560 TCTCCGTTGGCTGCTTGGCCATATTTGTCACGTGTGATCTCCCTCTGGTGTACAAAAAC 1619
QY 1587 ACAAGGAATACAAACCCCAATAGAAAATAGTCTCTGGGAATGTGGTCAGAAAGCAAGGCCCTGA 1646
Db 1620 ACAAGGAATACAAACCCCAATAGAAAATAGTCTCTGGGAATGTGGTCAGAAAGCAAGGCCCTGA 1679
QY 1647 GTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGAAACAGGAAAAGGATCCGC 1706
Db 1680 GTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGAAACAGGAAAAGGATCCGC 1739
QY 1707 TACTCAAAAACCAAGAAATTAAGGAGTTTCTTAAATTTTGACCTTGTTTCTGAAGCTCA 1766
Db 1740 TACTCAAAAACCAAGAAATTAAGGAGTTTCTTAAATTTTGACCTTGTTTCTGAAGCTCA 1799
QY 1767 CTTTTCAGTGCCCAATGATGTGAGATGTCTGGAGTGGCTATTAAACCTTTTTTCTTAAAG 1826
Db 1800 CTTTTCAGTGCCCAATGATGTGAGATGTCTGGAGTGGCTATTAAACCTTTTTTCTTAAAG 1859
QY 1827 ATTATTTGTTAAATAGATATTGTGGTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA 1886
Db 1860 ATTATTTGTTAAATAGATATTGTGGTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA 1919
QY 1887 TTTTAGAGATGGGGAGAGGATTATATCTGCAGGCAGCTTCAGCCATGTTGTGAAACTGAT 1946
Db 1920 TTTTAGAGATGGGGAGAGGATTATATCTGCAGGCAGCTTCAGCCATGTTGTGAAACTGAT 1979
QY 1947 AAAAGCAACTTAGCAAGGCTTCTTTTCATTATTTTATGTTTTCACCTTATAAAGTCTTAG 2006
Db 1980 AAAAGCAACTTAGCAAGGCTTCTTTTCATTATTTTATGTTTTCACCTTATAAAGTCTTAG 2039
QY 2007 GTAACCTAGTAGATAGAAAACACTGTGTCCCGAGAGTAAGGAGAGAAGCTACTATTGATTA 2066
Db 2040 GTAACCTAGTAGATAGAAAACACTGTGTCCCGAGAGTAAGGAGAGAAGCTACTATTGATTA 2099
QY 2067 GAGCCTAACCCAGGTTAACTGCAAGAAGAGCGGGATACCTTTCAGCTTTCATGTAACCTG 2126
Db 2100 GAGCCTAACCCAGGTTAACTGCAAGAAGAGCGGGATACCTTTCAGCTTTCATGTAACCTG 2159
QY 2127 TATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTTCCAGCTTAACCTGAATCCCA 2186
Db 2160 TATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTTCCAGCTTAACCTGAATCCCA 2219
QY 2187 CTTCAATACACACTCATGAACCTCCTGATGGAAACAATAACAGGCCCAAGCCTGTGGTATGA 2246
Db 2220 CTTCAATACACACTCATGAACCTCCTGATGGAAACAATAACAGGCCCAAGCCTGTGGTATGA 2279
QY 2247 TGTGCACACTTGCTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTTG 2306
Db 2280 TGTGCACACTTGCTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTTG 2339
QY 2307 GTGACAACTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATTCATTTATTCCTCA 2366



RESULT 5
HSNMB
LOCUS
DEFINITION H.sapiens NMB mRNA.
ACCESSION X76534
VERSION X76534.1 GI:666042
KEYWORDS NMB gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 2669)
AUTHORS Weterman,M.A., Ajubi,N., van Dinter,I.M., Degen,W.G., van
Muijen,G.N., Ruitter,D.J. and Bloemers,H.P.
TITLE nmb, a novel gene, is expressed in low-metastatic human melanoma
cell lines and xenografts
JOURNAL Int. J. Cancer 60 (1), 73-81 (1995)
MEDLINE 95113576
PUBMED 7814155
REFERENCE 2
AUTHORS Weterman,M.
TITLE Direct Submission
JOURNAL Submitted (03-DEC-1993) M. Weterman, University of Nijmegen, Dept
of Biochemistry, PO Box 9101, 6500 HB Nijmegen, NETHERLANDS
FEATURES
source
Location/Qualifiers
1. .2669
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NMB"
/cell_line="MV1"
/clone_lib="MV1"
1. .2669
/gene="NMB"
92. .1774
/gene="NMB"
/codon_start=1
/protein_id="CAA54044.1"
/db_xref="GI:666043"
/db_xref="GOA:Q14956"
/translation="MECLYFGLFLLAARPLDAAKRFDVLGNRPSAYMRHNQL
NGWSSDENDWNEKLYPVWKRGMRWKNSWKGRVQAVLTSDSPALVGSNITFAVNLIF
PRCKEDANGNI VYKNCRNEAGLSADPYVYNNWTAWSESDGNGTQSHNVFPDGK
PFPHPGWRRWNFIYVFHTLQYFQKLGRCRSVRVSVNTANVTLGPQLMEVTVYRRHGR
AYVPIAQVKDVYVVTQDIPVFTVMFQKNDNRNSDTEFLKDLPIMEDVLIHDPHSFLNY
STINYKWSFGDNTGLFVSTNHTVNHTYVNLNGTFSNLNLTVKAAAPGPCPPPPPPPSK
PTPSLGPAGDNPLBSRIPDENCOINRYGHFQATITIVEGILEVNI IQMTDVLMPVPW
PESSLIDFVVTCCQSIPTVECTIISDPTCEITQNTVCSPVDVDEMCLLTVRRTFNGSG
TYCVNLTLTGDDTSLALTSTLISVPDRDPASPLRMANSALISVGCLAI FVTIVISLLVYK

gene
CDS
1. .2669
/gene="NMB"
92. .1774
/gene="NMB"
/codon_start=1
/protein_id="CAA54044.1"
/db_xref="GI:666043"
/db_xref="GOA:Q14956"
/translation="MECLYFGLFLLAARPLDAAKRFDVLGNRPSAYMRHNQL
NGWSSDENDWNEKLYPVWKRGMRWKNSWKGRVQAVLTSDSPALVGSNITFAVNLIF
PRCKEDANGNI VYKNCRNEAGLSADPYVYNNWTAWSESDGNGTQSHNVFPDGK
PFPHPGWRRWNFIYVFHTLQYFQKLGRCRSVRVSVNTANVTLGPQLMEVTVYRRHGR
AYVPIAQVKDVYVVTQDIPVFTVMFQKNDNRNSDTEFLKDLPIMEDVLIHDPHSFLNY
STINYKWSFGDNTGLFVSTNHTVNHTYVNLNGTFSNLNLTVKAAAPGPCPPPPPPPSK
PTPSLGPAGDNPLBSRIPDENCOINRYGHFQATITIVEGILEVNI IQMTDVLMPVPW
PESSLIDFVVTCCQSIPTVECTIISDPTCEITQNTVCSPVDVDEMCLLTVRRTFNGSG
TYCVNLTLTGDDTSLALTSTLISVPDRDPASPLRMANSALISVGCLAI FVTIVISLLVYK

KHKEYNPIENSPGNVVRSKGLSVFLNRAKAVFFPGNQEKDPLLKNQEPFKGVs"	
mat_peptide	155. .1771
/gene="NMB"	
/product="unnamed"	
ORIGIN	
Query Match 95.7%; Score 2547.8; DB 9; Length 2669;	
Best Local Similarity 99.2%; Pred. No. 0;	
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;	
QY	28 AACCTTGGTGGCTCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCCCTGGGAT 87
Db	60 ACCTTGAGTGGCTCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCCCTGGGAT 119
QY	88 TTCTGCTCCTGCTGCAAGATTGCCACTTGTATGCCCAACAAACGATTTTCATGATGTCTGG 147
Db	120 TTCTGCTCCTGCTGCAAGATTGCCACTTGTATGCCCAACAAACGATTTTCATGATGTCTGG 179
QY	148 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGSGCTGCTTCTG 207
Db	180 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGSGCTGCTTCTG 239
QY	208 ATGAAAATGACTGGAATGAAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 267
Db	240 ATGAAAATGACTGGAATGAAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 299
QY	268 AAAACTCCTGGAAGGAGGCCGTGTGCAGGCGTCTGCACCATGACTACCCAGCCCTCG 327
Db	300 AAAACTCCTGGAAGGAGGCCGTGTGCAGGCGTCTGCAGGCGTCTGCACCATGACTACCCAGCCCTCG 359
QY	328 TGGGCTCAAATATAACATTTGCGGTGAACCTGATATCCCTAGATGCCAAAAGGAAGATG 387
Db	360 TGGGCTCAAATATAACATTTGCGGTGAACCTGATATCCCTAGATGCCAAAAGGAAGATG 419
QY	388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 447
Db	420 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 479
QY	448 CATATGTTTACAACCTGGACAGCATGGTCAGAGGACAGTGACGGGGAATAATGGCACCGGCC 507
Db	480 CATATGTTTACAACCTGGACAGCATGGTCAGAGGACAGTGACGGGGAATAATGGCACCGGCC 539
QY	508 AAAGCCCATATAACGCTTCCCTGATGGGAAACCTTTCTCACACCCCGGATGGAGAA 567
Db	540 AAAGCCCATATAACGCTTCCCTGATGGGAAACCTTTCTCACACCCCGGATGGAGAA 599
QY	568 GATGGAATTTTCATCTACGCTTCCACACACACTTGGTCAGTAATTCAGAAATTTGGGACGAT 627
Db	600 GATGGAATTTTCATCTACGCTTCCACACACACTTGGTCAGTAATTTCCAGAAATTTGGGACGAT 659
QY	628 GTTCAGTGAGAGTTTCTGTGAACACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAG 687
Db	660 GTTCAGTGAGAGTTTCTGTGAACACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAG 719
QY	688 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 747
Db	720 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 779
QY	748 ACGTGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAAGAACGATCGAAATT 807
Db	780 ACGTGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAAGAACGATCGAAATT 839
QY	808 CATCCGACGAAACCTTCC - CAAAGATCTCCCATTTATGTTTGTATGCTCTGATTCATGATC 866
Db	840 CATCCGACGAAACCTTCC - CAAAGATCTCCCATTTATGTTTGTATGCTCTGATTCATGATC 899
QY	867 CTAGCCACTTCCCTCAATATTCTACCAATTAACCTGAGTGGAGCTTCGGGGATAATACTG 926
Db	900 CTAGCCACTTCCCTCAATATTCTACCAATTAACCTGAGTGGAGCTTCGGGGATAATACTG 959
QY	927 GCCTGTTTGTTCACCAATCATACTGTGAATCACACGATGTGCTCAATGGAACTTCA 986
Db	960 GCCTGTTTGTTCACCAATCATACTGTGAATCACACGATGTGCTCAATGGAACTTCA 1019

QY	987 GCCTTAAACCTCACTGTGAAAGCTGCAGCACCAAGGACCTTGTGCGCCACCGCCACCAACAC 1046
Db	1020 GCCTTAAACCTCACTGTGAAAGCTGCAGCACCAAGGACCTTGTGCGCCACCGCCACCAACAC 1079
QY	1047 CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGTGTTGACAAACCCCTCGGAGCTGA 1106
Db	1080 CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGTGTTGACAAACCCCTCGGAGCTGA 1139
QY	1107 GTAGGATTCCTGATGAAAACTGCCAGATTAACAGATATGGCCACTTTCAAGCCACCATCA 1166
Db	1140 GTAGGATTCCTGATGAAAACTGCCAGATTAACAGATATGGCCACTTTCAAGCCACCATCA 1199
QY	1167 CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGG 1226
Db	1200 CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGG 1259
QY	1227 TGCCATGGCCTGAAAAGCTCCCTAATAGACTTTGTGTCGTCGTCCTGCAAGGGAGCATTTCCCA 1286
Db	1260 TGCCATGGCCTGAAAAGCTCCCTAATAGACTTTGTGTCGTCGTCCTGCAAGGGAGCATTTCCCA 1319
QY	1287 CGGAGGTCCTGATACCATCATTTCTGACCCACCTGCGAGATCAACCAGAACACAGTCTGCA 1346
Db	1320 CGGAGGTCCTGATACCATCATTTCTGACCCACCTGCGAGATCAACCAGAACACAGTCTGCA 1379
QY	1347 GCCCTGTGGATGTGATGATGTGTCGTGACTGTGAGACGAACCTTCAATGGGTCTG 1406
Db	1380 GCCCTGTGGATGTGATGATGTGTCGTGACTGTGAGACGAACCTTCAATGGGTCTG 1439
QY	1407 GGACGTACTGTGTGAAACCTCACCCCTGGGGGATGACACAAGCCTGGCTCTCAGGACACCC 1466
Db	1440 GGACGTACTGTGTGAAACCTCACCCCTGGGGGATGACACAAGCCTGGCTCTCAGGACACCC 1499
QY	1467 TGATTTCTGTTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCTTGA 1526
Db	1500 TGATTTCTGTTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCTTGA 1559
QY	1527 TCTCCGTTGGCTGTGGCCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAAC 1586
Db	1560 TCTCCGTTGGCTGTGGCCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAAC 1619
QY	1587 ACAAGGAATACAAACCCATAGAAAAATAGTCCTGGGAATGTGGTCAGAAAGCAAAGGCTGCA 1646
Db	1620 ACAAGGAATACAAACCCATAGAAAAATAGTCCTGGGAATGTGGTCAGAAAGCAAAGGCTGCA 1679
QY	1647 GTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGAAACCCAGGAAAGGATCCGC 1706
Db	1680 GTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGAAACCCAGGAAAGGATCCGC 1739
QY	1707 TACTCAAAAAACCAAGAAATTTAAAGGAGTTTCTTAAATTTTCGACCTTGTCTTGAAGCTCA 1766
Db	1740 TACTCAAAAAACCAAGAAATTTAAAGGAGTTTCTTAAATTTTCGACCTTGTCTTGAAGCTCA 1799
QY	1767 CTTTTCAGTGCCATGATGTGAGATGTGCTGGAGTGGCTATTAACCTTTTTTTCCTAAAG 1826
Db	1800 CTTTTCAGTGCCATGATGTGAGATGTGCTGGAGTGGCTATTAACCTTTTTTTCCTAAAG 1859
QY	1827 ATTTATGTTAAATAGATATTTGGTTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA 1886
Db	1860 ATTTATGTTAAATAGATATTTGGTTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA 1919
QY	1887 TTTTATGATGGGAGAGGGATTATATCTGCAGGCAGCTTCAGCCATGTTGTGAAACTGAT 1946
Db	1920 TTTTATGATGGGAGAGGGATTATATCTGCAGGCAGCTTCAGCCATGTTGTGAAACTGAT 1979
QY	1947 AAAAGCAACTTAGCAAGGCTTCTTTTCATTTATTTTTTATGTTTCACTTATAAAGTCTTAG 2006
Db	1980 AAAAGCAACTTAGCAAGGCTTCTTTTCATTTATTTTTTATGTTTCACTTATAAAGTCTTAG 2039
QY	2007 GTAACCTAGTAGGATAGAAACACATGTGTCCCGAGAGTAAGGAGAGAACTACTATTGATTA 2066
Db	2040 GTAACCTAGTAGGATAGAAACACATGTGTCCCGAGAGTAAGGAGAGAACTACTATTGATTA 2099

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 10, 2004, 10:01:28 ; Search time 40.5 Seconds
(without alignments)
6842.404 Million cell updates/sec

Title: US-10-039-272-1
Perfect score: 4850
Sequence: 1 cggcacgagggccagagga.....aaaaaaaaaaaaaaaaaaaa 2661

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10039272/runat_07092004_144222_21619/app query.fasta_1.2823
-DB=SwissProt_42 -QFMT=fastan -SUFFIX=rsp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10039272@cgn 1 1 48 @runat_07092004_144222_21619 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

		SUMMARIES			
Result No.	Score	Query Match	Length	DB ID	Description
1	2999	61.8	560	1	NMB_HUMAN
2	1563.5	32.2	559	1	QNR_COTJA
3	635	13.1	626	1	PM17_MOUSE
4	617	12.7	762	1	P115_CHICK
5	593	12.2	661	1	PM17_HUMAN
6	436.5	9.0	491	1	PM17_BOVIN
7	129	2.7	926	1	PTN4_HUMAN
8	116.5	2.4	555	1	ODP2_RAT
9	115.5	2.4	1093	1	SM5B_MOUSE
10	111	2.3	525	1	HRG_HUMAN
11	111	2.3	5038	1	PCLO_MOUSE
12	110.5	2.3	1874	1	POLR_KYMVJ
13	110	2.3	1225	1	CTD2_HUMAN
14	109.5	2.3	5703	1	MU5B_HUMAN
15	106.5	2.2	427	1	IRF3_HUMAN
16	106	2.2	1537	1	FLO1_YEAST
17	104	2.1	435	1	GAT3_XENLA
18	104	2.1	4303	1	PKD1_HUMAN

19	103.5	2.1	5147	1	PCLO_HUMAN	Q9y6v0	homo sapien
20	103	2.1	626	1	GPBA_HUMAN	P07359	homo sapien
21	103	2.1	921	1	RB_MOUSE	P13405	mus musculus
22	103	2.1	1009	1	YE68_METJA	Q58863	methanococc
23	102.5	2.1	579	1	SOC6_MOUSE	Q8vhq2	mus musculus
24	102.5	2.1	979	1	RFX1_HUMAN	P22670	homo sapien
25	102	2.1	1213	1	FMN_CHICK	Q05858	gallus gall
26	101	2.1	2090	1	N214_HUMAN	P35658	homo sapien
27	100	2.1	1068	1	DAM2_HUMAN	Q86t65	homo sapien
28	99.5	2.1	1587	1	LMG3_HUMAN	Q9y6n6	homo sapien
29	99	2.0	450	1	AT10_MOUSE	P58459	mus musculus
30	99	2.1	499	1	CHIB_SERMA	P11797	serratia ma
31	98.5	2.0	594	1	RYK_MOUSE	Q01887	mus musculus
32	98	2.0	1059	1	CAPU_DROME	Q24120	drosophila
33	98	2.0	4548	1	APOA_HUMAN	P08519	homo sapien
34	98	2.0	5120	1	PCLO_CHICK	Q9pu36	gallus gall
35	97.5	2.0	785	1	CAD7_HUMAN	Q9ulb5	homo sapien
36	97.5	2.0	1858	1	P3K2_DICDI	P54674	dictyosteli
37	97	2.0	340	1	CYSP_SCHMA	P25792	schistosoma
38	97	2.0	673	1	REP_ECOLI	P09980	escherichia
39	96.5	2.0	488	1	BTBI_MOUSE	P58544	mus musculus
40	96	2.0	377	1	AR41_SCHPO	P78774	schizosacch
41	96	2.0	512	1	PER1_VOLCA	P81131	volvox cart
42	95.5	2.0	671	1	ENV_FENV1	P31791	feline endo
43	95.5	2.0	1161	1	DAN4_YEAST	P47179	saccharomyc
44	95.5	2.0	1247	1	SAH1_HUMAN	O94885	homo sapien
45	95.5	2.0	1620	1	HED1_HUMAN	Q9ult8	homo sapien

ALIGNMENTS

RESULT 1
NMB_HUMAN
ID NMB_HUMAN STANDARD; PRT; 560 AA.
AC Q14956;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, last sequence update)
DT 15-MAR-2004 (Rel. 43, last annotation update)
DE Putative transmembrane protein NMB precursor (Transmembrane glycoprotein HGFIN).
DE GPNMB OR NMB OR HGFIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Melanoma;
RX MEDLINE=95113576; PubMed=7814155;
RA Weterman M.A.J., Ajubi N., van Dinter I.M.R., Degen W.G.J., van Muijen G.N.P., Ruiter D.J., Bloemers H.P.J.;
RT "NMB, a novel gene, is expressed in low-metastatic human melanoma cell lines and xenografts";
RL Int. J. Cancer 60:73-81(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Peripheral blood;
RX MEDLINE=22498106; PubMed=12609765;
RA Bandari P.S., Qian J., Yehia G., Joshi D.D., Maloof P.B., Potian J., Oh H.S., Gascon P., Harrison J.S., Rameshwar P.;
RT "Hematopoietic growth factor inducible neurokinin-1 type: a transmembrane protein that is similar to neurokinin 1 interacts with substance P";
RL Regul. Pept. 111:169-178(2003).
CC -!- FUNCTION: Could be a melanogenic enzyme (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: Not restricted to the melanocytic lineage.
CC -!- DEVELOPMENTAL STAGE: Expression in poorly metastatic melanoma cell lines; no expression in highly metastatic melanoma cell lines.
CC -!- SIMILARITY: BELONGS TO THE PMEL-17/NMB FAMILY.
CC -!- SIMILARITY: Contains 1 PKD domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; X76534; CAA54044.1; --
DR EMBL; AF322309; AAG42839.1; --
DR PIR; I38065; I38065.
DR Genew; HGNC:4462; GPNMB.
DR MIM; 604368; --
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR InterPro; IPR00601; PKD.
DR Pfam; PF00801; PKD; 1.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS50093; PKD; 1.
KW Signal; Transmembrane; Glycoprotein; Polymorphism.
FT SIGNAL 1 21
FT CHAIN 22 560
FT DOMAIN 22 486
FT TRANSMEM 487 507
FT DOMAIN 508 560
FT DOMAIN 240 327
FT DOMAIN 320 332
FT CARBOHYD 93 93
FT CARBOHYD 134 134
FT CARBOHYD 146 146
FT CARBOHYD 200 200
FT CARBOHYD 249 249
FT CARBOHYD 275 275
FT CARBOHYD 296 296
FT CARBOHYD 300 300
FT CARBOHYD 306 306
FT CARBOHYD 312 312
FT CARBOHYD 447 447
FT CARBOHYD 455 455
FT VARIANT 195 195
FT VARIANT 197 197
FT VARIANT 197 197
FT SEQUENCE 560 AA; 62643 MW; 570035B48CCE3ECC CRC64;
Alignment Scores:
Pred. No.: 3.37e-248 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 1 Gaps: 0

US-10-039-272-1 (1-2661) x NMB_HUMAN (1-560)
QY 60 ATGGAATGCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTTCATGATGTCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAATGGCTGGTCTTCTGATGAAATGACTGGAAATGAAATACTTACCCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGlnLysLeuTyrPro 60
QY 240 GTGTGGAAGCGGGGAGACATGAGGTGGAAATACTCCTGGAAGGAGCGCGTGTGAGGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80
QY 300 GTCCCTGACCACTGACTCACCAGCCCTCGTGGGTCAAATATAACATTTCGGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100

QY 360 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGCTGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGGAATAATGSCACCGGCCAAAGCCATACATAACGTCCTTCCCTGATGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CTTTTCCTCACCACCCCGGATGGAGAGATGGAATTTTCATCTACGTCCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCACTATTCCAGAAATTTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGCCTCAACTCATGGAAGTACTGTCTACAGAGACATGGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTGAACAGATCAGATTCCTGTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTGTGCTGATTTCATGATCCTTACCCACTTCTCTCAATTATTTACCATTAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGTTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATACTCTGTAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCACCACCCAGACCTTCAAAACCCACCCCTTCTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProProProProProProSerLeuGly 340
QY 1079 CCTGTGTTGACAAACCCCTGGAGCTGAGTAGGATTCTCTGATGAAACTGCCAGATTAAAC 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTTCAAGCCACCACCATCACAAATTGTAGAGGGAATCTTAGAGGTTAACATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGTCCTGATGCCGTGCCATGGCCTGAAAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400
QY 1259 GTCTGACCTGCCAAGGGAGCATTCCTGATGCCGTGCCATGGCCTGAAAGCTCCCTAATAGACTTT 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TCGAGATCACCCAGAACACAGTCTGTCAGCCCTGTGGATGTGGATGTGGATGTGTCTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGACGAAACCTTCAATGGTCTGGGACGCTACTGTGTGAACCTCACCCCTGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460

QY 1439 GACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCCTGACAGAGACCCAGCCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGGATGGCAAAACAGTGCCTGATCTCCGTGGCTGGCTTGGCCATATTGTCACT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTTGGTGTTACAAAAACACAAGGAATACACCCCAATAGAAAATAGTCTT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGGATGTGGTCAGAAGCAAGGCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTC 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGAAACAGGAAAGATCCGCTACTCAAAAACCAAGAATTAAAGGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 2
QNR_COTJA
ID QNR_COTJA STANDARD; PRT; 559 AA.
AC Q90372;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE QNR-71 protein precursor.
GN QNR-71.
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=96283633; PubMed=8670835;
RA Turque N., Denhez F., Martin P., Planque N., Bailly M., Begue A.,
RA Stehelin D., Saule S.;
RT "Characterization of a new melanocyte-specific gene (QNR-71)
RT expressed in v-myc-transformed quail neuroretina."
RL BMBO J. 15:3338-3350(1996).
RN [2]
RP REVISIONS.
RC TISSUE=Retina;
RA Saule S.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Could be involved in melanogenesis.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: Melanocyte-specific, restricted to the
CC pigmented layer of the retina and the epidermis.
CC -!- DEVELOPMENTAL STAGE: Transcriptionally regulated by MYC in the
CC transdifferentiation of embryo pigmented epithelial cells.
CC Expression precedes melanization.
CC -!- SIMILARITY: BELONGS TO THE PMEL-17/NMB FAMILY.
CC -!- SIMILARITY: Contains 1 PKD domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X94144; CAA63859.1; -.
DR InterPro; IPR000601; PKD.
DR Pfam; PF00801; PKD; 1.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS50093; PKD; 1.
KW Signal; Transmembrane; Glycoprotein.
FT SIGNAL 1 22 POTENTIAL.

FT CHAIN 23 559 QNR-71 PROTEIN.
FT DOMAIN 23 487 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 488 508 POTENTIAL.
FT DOMAIN 509 559 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 239 326 PKD.
FT CARBOHYD 92 92 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 149 149 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 311 311 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 559 AA; 61713 MW; 81E2793BD64A47DD CRC64;

Alignment Scores:
Pred. No.: 1.89e-125 Length: 559
Score: 1563.50 Matches: 297
Percent Similarity: 68.83% Conservative: 85
Best Local Similarity: 53.51% Mismatches: 158
Query Match: 32.24% Indels: 16
DB: 1 Gaps: 5

US-10-039-272-1 (1-2661) x QNR_COTJA (1-559)

QY 81 CTGGGATTTCTGCTCCTGGTGCAAGATTGCCACTTGATGCCGCCAAACGATTTCATGAT 140
Db 8 LeuAlaLeuLeuLeuProAlaGluAlaValLeuCysAlaAlaMetArgPheGlnAsp 27

QY 141 GTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTAATGGCTGG 200
Db 28 ValLeuSerAsnGlyArg---ThrAlaProValThrAsnHisLysLysIleGlnGlyTrp 46

QY 201 TCTTCTGATGAAAATGACTGGAATGAAAACCTCTACCCAGTGTGGAAGCGGGAGACATG 260
Db 47 SerSerAspGlnAsnLysTrpAsnGluLysLeuTyrProPheTrpGluAspAsnAspPro 66

QY 261 AGGTGGAAAACTCCTGGAAGGGAGGCCGTGTGTCAGGGCGTCTGACGAGTCACTACCA 320
Db 67 ArgTrpLysAspCysTrpLysGlyGlyLysValThrThrLysLeuValThrAspSerPro 86

QY 321 GCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAG 380
Db 87 AlaLeuValGlySerAsnValThrPheValValThrLeuGlnPheProLysCysGlnLys 106

QY 381 GAAGATGCCAATGGCAACATAGTCTATGAGAAGAAGAACTGCAGAAATGAGGCTGTTATCT 440
Db 107 GluAspAspAspGlyAsnIleIleTyrGlnArgAsnCysThrProAspSerProAlaAla 126

QY 441 GCTGATCCATATGTTTACAACGTGACAGCATGGTCAGAGGACAGTGCAGGGGAAAATGGC 500
Db 127 GlnAspGlnTyrValTyrAsnTrpThrGluTrpIleAspAsnCysGlyTrpGluAsnCys 146

QY 501 ACCGGCCAAAAGCCATCATAACGTCTTCCCTGATGGGAAACCTTTTCTCACCCACCCCGA 560
Db 147 ThrSerAsnHisSerHisAsnValPheProAspGlyLysProPheProHisTyrProGly 166

QY 561 TGGAGAAAGATGGAATTTTCATCTACGTCTTCCACACACACITGGTCAGTATTCCAGAAATTG 620
Db 167 TrpArgArgArgAsnPheValTyrLeuPheHisThrValGlyGlnTyrTyrGlnThrIle 186

QY 621 GGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCCTCACTC 680
Db 187 GlyArgSerSerAlaAsnPheSerValAsnThrAlaAsnIleThrLeuGlyLysHisIle 206

QY 681 ATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTAA 740
Db 207 MetAlaValSerIleTyrArgArgGlyHisSerThrTyrValProIleAlaArgAlaSer 226

QY 741 GATGTGTACGTGGTAAACAGATCAGATTCTCTGTGTTTGTGACTATGTTCCAGAAGACCAT 800

Db 227 ThrThrValValThrAspGlyPheProIleLeuValSerMetSerGlnHisAsp 246
QY 801 CGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCATATGTTTGTATGCTGATT 859
Db 247 ArgAsnIleSerAspSerIlePheIleLysAspSerProIleThrPheAspValLysIle 266
QY 860 CATGATCCTAGCCACTTCTCAATATTCTACCATTAACAAGTGGAGCTTCGGGGAT 919
Db 267 HisAspProSerTyrTyrLeuAsnAspSerAlaIleSerTyrLysTrpAsnPheGlyAsp 286
QY 920 AATACTGGCCTGTTTGTTCACCAATCATACTGTGAATCACACGATGTGTCTCAATGGA 979
Db 287 GlySerGlyLeuPheValGluSerGlyAlaThrThrSerHisThrPheSerLeuGlnGly 306
QY 980 ACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACGAGACCTTGTCCGCCACCGCCA 1039
Db 307 AsnPheThrLeuAsnLeuThrValGlnAlaIleIleProValProCys-----LysPro 324
QY 1040 CCACCACCCAGACCTTCAAAACCCACCCCTTCTTTA----- 1075
Db 325 ValThrProThrProSerLeuProThrProAlaValThrThrAspAlaSerSerAsnSer 344
QY 1076 GGACCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACTGCCAGATT 1135
Db 345 AspProSerAlaProAsnGluMetAlaGluAspAsn--ProAspGlyGlyCysHisIle 363
QY 1136 AACAGATATGGCCACTTTCAGCCACCATCACAATTTAGAGGGATCTTAGAGGTTAAC 1195
Db 364 TyrArgTyrGlyTyrTyrThrAlaGlyIleThrIleValGluGlyIleLeuGluValAsn 383
QY 1196 ATCATCCAGATGACAGACGTCCTGATGCGCGTGCCATGGCCTGAAAGCTCCCTAATAGAC 1255
Db 384 IleIleGlnMetThrSerIleGlnMetThrGluSerGlnAlaGluAsnProLeuValAsp 403
QY 1256 TTTGTCGTGACTGCCAAGGAGGACTTCCACGGAGGCTGTGACCATCATTTCTGACCCC 1315
Db 404 PheValValThrCysGlnGlySerPheProThrAspValCysThrAlaValSerAspPro 423
QY 1316 ACCTGCGAGATCACCCAGAAACACAGTCTGAGCCCTGTGGATGTGGATGAGATGTGCTG 1375
Db 424 ThrCysGlnValSerGlnGlyMetValCysAspProValValValThrAspGluCysVal 443
QY 1376 CTGACTGTGAGACGAACCTTCAATGGGTCTGGAGCTACTGTGTGAACCTCACCCCTGGGG 1435
Db 444 LeuThrIleArgAlaPheAspGluProGlyThrTyrCysIleAsnIleThrLeuGly 463
QY 1436 GATGACACAAGCCTGCTCTCACGAGCACCCCTGATTCTGTCTCTGACAGAGACCCAGCC 1495
Db 464 AspAspThrSerGlnAlaLeuAlaSerAlaLeuIleSerVal-----AsnGlyGly 480
QY 1496 TCGCCTTTAAGGATGGCAAAACAGTGCCTGATCTCGTGGCTGTGCTTGGCCATATTGTC 1555
Db 481 SerSerSerGlyThrThrLysGlyValPheIlePheLeuGlyLeuLeuAlaValPheGly 500
QY 1556 ACTGTGATCTCCCTCTTGGTGTAACAAAACACAGGAATACACCCCAATAGAAAATAGT 1615
Db 501 AlaIleGlyAlaPheValLeuTyrLysArgTyrLysGlnTyrLysProIleGluArgSer 520
QY 1616 CCTGGGAATGTGGTCAGAAAGCAAGGCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTG 1675
Db 521 AlaGlyGlnAlaGluAsnGlnGluGlyLeuSerAlaTyrValSerAsnPheLysAlaPhe 540
QY 1676 TTCTTCCCGGAAACCCAGGAAAAGGATCCGCTACTCAAAAACCAA 1720
Db 541 PhePheProLysSerThrGluArgAsnProLeuLeuLysSerLys 555

RESULT 3
PM17_MOUSE
ID PM17_MOUSE STANDARD; PRT; 626 AA.
AC Q60696;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Melanocyte protein Pmel 17 precursor (Silver locus protein).
GN SILV OR PMEL17 OR D10H12553E OR SI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
STRAIN=C57BL/6; TISSUE=Skin;
MEDLINE=95175358; PubMed=7870580;
RA Kwon B.S., Halaban R., Ponnazhagan S., Kim K., Chintamaneni C.,
Bennett D., Pickard R.T.;
RT "Mouse silver mutation is caused by a single base insertion in the
RT putative cytoplasmic domain of Pmel 17.";
RL Nucleic Acids Res. 23:154-158(1995).
CC -!- FUNCTION: Could be a melanogenic enzyme.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- TISSUE SPECIFICITY: Preferentially expressed in melanocytes.
CC -!- DISEASE: DEFECTS IN SILV ARE THE CAUSE OF THE SILVER COAT COLOR
CC WHICH SEEMS TO BE DUE TO PREMATURE DEATH OF PIGMENT CELLS DURING
CC THE HAIR CYCLE.
CC -!- SIMILARITY: BELONGS TO THE PMEL-17/NMB FAMILY.
CC -!- SIMILARITY: Contains 1 PKD domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; U14133; AAA69538.1; -.
DR PIR; S53871; S53871.
DR MGD; MGI:98301; Si.
DR InterPro; IPR000601; PKD.
DR Pfam; PF00801; PKD; 1.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS50093; PKD; 1.
KW Transmembrane; Glycoprotein; Signal; Melanin biosynthesis; Repeat;
KW Disease mutation.
KW SIGNAL 1 24
FT CHAIN 25 626
FT DOMAIN 25 562
FT TRANSMEM 563 583
FT DOMAIN 584 626
FT DOMAIN 255 292
FT DOMAIN 315 411
FT REPEAT 315 327
FT REPEAT 328 340
FT REPEAT 341 353
FT REPEAT 354 366
FT REPEAT 367 379
FT REPEAT 380 392
FT REPEAT 393 411
FT CARBOHYD 81 81
FT CARBOHYD 106 106
FT CARBOHYD 111 111
FT CARBOHYD 535 535
FT VARIANT 170 170
FT VARIANT 175 175
FT VARIANT 373 373
FT VARIANT 471 471
FT VARIANT 603 626
SQ SEQUENCE 626 AA; 65980 MW; 7AB941D2E3FB1044 CRC64;
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
N-LINKED (GLCNAC. .) (POTENTIAL).
S -> L (IN SILVER).
R -> G (IN SILVER).
D -> N (IN SILVER).
F -> S (IN SILVER).
AAPASGLRARGLGENSEPLLSGQQV -> SSASLSRRPRPW
RKQAPQWTAGLIILKAPWISWG (IN SILVER).
7AB941D2E3FB1044 CRC64;

Alignment Scores:
Pred. No.: 4.88e-46 Length: 626
Score: 635.00 Matches: 184
Percent Similarity: 41.45% Conservative: 97
Best Local Similarity: 27.14% Mismatches: 207

Query Match:	13.09%	Indels:	191
DB:	1	Gaps:	22
US-10-039-272-1 (1-2661) x PM17_MOUSE (1-626)			
QY	78	TTCTTGGGATTTC	131
Db	9	PheLeuProValLeuValLeuSerAlaLeuLeuAlaValGlyAlaLeuGluGlySerArg	28
QY	132	TTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTA	191
Db	29	AsnGlnAspTrpLeuGlyValProArgGlnLeuVal	40
QY	192	AATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTTACCCAGTGTGGAGCGG	251
Db	41	-----ThrLysThrTrpAsnArgGlnLeuTrpProGluTrpThr---	53
QY	252	GGAGACATGAGGTGGAAAACTCCTGGAAGGGAGGCGGTGTCAGGCGTCTCAGCACAGT	311
Db	54	---GluValGlnGlySerAsnCysTrpArgGlyGlyGlnValSerLeuArgValIleAsn	72
QY	312	GACTCACCGCCCTCGTGGCTCAAATATAACATTTGCGGTGAACCTGATATCCCTAGA	371
Db	73	AspGlyProThrLeuValGlyAlaAsnAlaSerPheSerIleAlaLeuHisPheProGly	92
QY	372	TGCCAAAAGGAATGCCAATGGCAACATAGTCTATGAGAAAGAAC---	428
Db	93	SerGlnLysValLeuProAspGlyGlnValIleTrpAlaAsnThrIleIleAsnGly	112
QY	429	GCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAGGACAGTGAC	488
Db	113	SerGlnValTrpGlyGlyGlnProValTyPro-----GlnGluProAspAsp	128
QY	489	GGGAAAAATGGCACCGGCCCAAGCCATCATAAACGTCTTCCCTGATGGGAAACCTTTTCCT	548
Db	129	Ala-----CysValPheProAspGlyGlyProCysPro	139
QY	549	CACCACCCCGGATGGAGAAGATGGAATTTCACTACGTCTTCCACACACTTGGTCAGTAT	608
Db	140	SerGlyProLysProProLysArgSerPheValTyValTrpLysThrTrpGlyLysTy	159
QY	609	TTCCAGAAATTTGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAATGTGACACTT	668
Db	160	TrpGlnValLeuGlyGlyProValSerArgSerSerIleAlaThrArgHisAlaLysLeu	179
QY	669	GGGCCTCAACTCATGGAAGTACTGTCTACAGAAGACATGGA---	725
Db	180	GlyThrHisThrMetGluValThrValTyHisArgArgGlySerGlnSerTyValPro	199
QY	726	ATCGCAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCTCTGTGTTTGTGACTATG	785
Db	200	LeuAlaHisAlaSerSerThrPheThrIleThrAspGlnValProPheSerValSerVal	219
QY	786	TTCCAGAA-----GAACGATCGAAATTCATCCGACGAAACCTTCCCAA	829
Db	220	SerGlnLeuGlnAlaLeuAspGlyGluThrLysHisPheLeuArgAsnHis-----	236
QY	830	GATCTCCCATATATGTTTGATGTCCTGATTTCATGATCCTAGCCACTTCCTCAATTATCT	889
Db	237	-----ProLeuIlePheAlaLeuGlnLeuHisAspProSerGlyTyTrpLeuAlaGluAla	254
QY	890	ACCATTAACTACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTGTTTCCACCAATCAT	949
Db	255	AspLeuSerTyTrpTrpAspPheGlyAspGlyThrGlyThrLeuIleSerArgAlaLeu	274
QY	950	ACTGTGAATCACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCT	1009
Db	275	AspValThrHisThrTyTrpLeuGluSerGlySerValThrAlaGlnValValLeuGlnAla	294
QY	1010	GCA-----	1012
Db	295	AlaIleProLeuValSerCysGlySerSerProValProGlyThrThrAspGlyTyMet	314

QY	1012	-----	1012
Db	315	ProThrAlaGluAlaProGlyThrThrSerArgGlnGlyThrThrLysValValGly	334
QY	1013	---GCACCAGGACCTTGTCCGCCACCGCCACCA---	1042
Db	335	ThrThrProGlyGlnMetProThrThrGlnProSerGlyThrThrValValGlnMetPro	354
QY	1042	-----	1042
Db	355	ThrThrGluValThrAlaThrThrSerGluGlnMetLeuThrSerAlaValIleAspThr	374
QY	1043	-----CCACCCAGACCTTCA---	1057
Db	375	ThrLeuAlaGluValSerThrThrGluGlyThrGlyThrThrProThrArgProSerGly	394
QY	1058	-----AAACCCACCCCTTCTTTAGGACCTGCTGGTGACAACCCCTG-----	1099
Db	395	ThrThrValAlaGlnAlaThrThrThrGluGlyProAspAla---SerProLeuLeuPro	413
QY	1100	-----GAGCTGAGTAGGATTCTCTGATGAA---	1123
Db	414	ThrGlnSerSerThrGlySerIleSerProLeuLeuAspAspThrAspThrIleMetLeu	433
QY	1124	-----AACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACC	1162
Db	434	ValLysArgGlnValProLeuAspCysValLeuTyArgTyArgTyGlySerPheSerLeuAla	453
QY	1163	ATCACAATTGTAGAGGGAATCTTAGAGGTTAACATCATCCAGATGACAGACGTCTCTGATG	1222
Db	454	LeuAspIleValGlnGly-----IleGluSerAlaGluIleLeuGln	467
QY	1223	CCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTCTGTCGTCCTGCCAAAGGAGCAT	1282
Db	468	AlaValProPheSerGluGlyAspAlaPheGluLeuThrValSerCysGlnGlyGlyLeu	487
QY	1283	CCCACGGAGTCTGTACCATCATTTCTGACCCACCTGCGAGATCACCCAGAACACAGTC	1342
Db	488	ProLysGluAlaCysMetAspIleSerSerProGlyCysGlnProProAlaGlnArgLeu	507
QY	1343	TGCAGCCCTGTGGATGTGGATGAGATGTCTGTCTGCTGACTGTGAGACGAAACCTTC---	1399
Db	508	CysGlnSerValProProSerProAspCysGlnLeuValLeuHisGlnValLeuLysGly	527
QY	1400	GGGTCTGGGACGTACTGTGTGAACCTCACCTCGGGGATGACACAGCCTGGCTCTCACG	1459
Db	528	GlySerGlyThrTyCysLeuAsnValSerLeuAlaAspAlaAsnSerLeuAlaValAla	547
QY	1460	AGCACCCCTGATTTCTGTTCTTCACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACAGT	1519
Db	548	SerThrGlnLeuValValProGlyGlnAspGlyGly-----LeuGlyGln	562
QY	1520	GCCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATC-----TCCCTCTTG	1573
Db	563	AlaProLeuLeuValGlyIleLeuLeuValLeuValAlaValValLeuAlaSerLeuIle	582
QY	1574	GTG-----TACAAAAAACACAAAGGAATACAAC-----CCAAATAGAA	1609
Db	583	LeuGlyIleAspLeuArgSerArgAlaGlnPheProLysCysHisMetValAlaLeuThr	602
QY	1610	AATAGTCTCTGGGAATGTGGTCAGAAAGCAAGGCCCTGAGTGTCTTTCTCAACCCGTGCAAAA	1669
Db	603	AlaAlaProAlaSerGlyLeuArgAlaArgGlyLeuGly-----	615
QY	1670	GCCGTGTTCTTCCCGGAAACCCAGGAAAGGATCCGCTACTCAAAAAACCAAGAA	1723
Db	616	-----GluAsnSerProLeuLeuSerGlyGlnGln	625
RESULT 4			
P115_CHICK			
ID	P115	CHICK	
AC	Q98917	STANDARD;	PRT; 762 AA.
DT	15-JUL-1998	(Rel. 36, Created)	

DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 115 kDa melanosomal matrix protein precursor.
GN MNP115.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=White leghorn; TISSUE=Retinal pigment epithelium;
RX MEDLINE=92020667; PubMed=1924173;
RA Mochii M., Agata K., Eguchi G.;
RT "Complete sequence and expression of a cDNA encoding a chicken
RT 115-kDa melanosomal matrix protein.";
RL Pigment Cell Res. 4:41-47(1991).
RN [2]
RP CHARACTERIZATION.
RC STRAIN=White leghorn; TISSUE=Retinal pigment epithelium;
RX MEDLINE=88311098; PubMed=3409326;
RA Mochii M., Agata K., Kobayashi H., Yamamoto T.S., Eguchi G.;
RT "Expression of gene coding for a melanosomal matrix protein
RT transcriptionally regulated in the transdifferentiation of chick
RT embryo pigmented epithelial cells.";
RL Cell Differ. 24:67-74(1988).
CC -!- FUNCTION: Might be required for polymerization of melanin onto the
CC core structure of melanosomes with enzymic function of tyrosinase.
CC -!- SUBCELLULAR LOCATION: ON THE FIBROUS MATRIX STRUCTURE OF THE
CC PREMELANOSOME.
CC -!- TISSUE SPECIFICITY: Specific to pigmented epithelial cells and
CC melanocytes. Not expressed in lens, neural retina, brain, heart,
CC gizzard or liver.
CC -!- DEVELOPMENTAL STAGE: Expressed during the redifferentiation of
CC pigmented epithelial cells (PEC).
CC -!- PTM: Glycosylated.
CC -!- SIMILARITY: BELONGS TO THE PMEL-17/NMB FAMILY.
CC -!- SIMILARITY: Contains 1 PKD domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D88348; BAA13589.1; -.
DR InterPro; IPR000601; PKD.
DR Pfam; PF00801; PKD; 1.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS50093; PKD; 1.
KW Signal; Glycoprotein; Repeat.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 762 115 KDA MELANOSOMAL MATRIX PROTEIN.
FT DOMAIN 223 323 PKD.
FT DOMAIN 441 532 4 X 20-24 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 441 464 1.
FT REPEAT 465 488 2.
FT REPEAT 489 508 3.
FT REPEAT 509 532 4.
FT CARBOHYD 111 111 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 346 346 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 651 651 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 659 659 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 762 AA; 77356 MW; 172C8DDB4FDC67C6 CRC64;

Alignment Scores: 1.81e-44 Length: 762
Pred. No.: 617.00 Matches: 171
Score: 35.60% Conservative: 80
Percent Similarity: 24.26% Mismatches: 186
Best Local Similarity:

Query Match:	12.72%	Indels:	269
DB:	1	Gaps:	15
US-10-039-272-1 (1-2661) x P115_CHICK (1-762)			
QY 90	CTGCTCCTGGTGCAGAGATTGCCACTTGTATGCGCGCAACGATTTCATGATGTGTGGGC	149	
DB 8	ValLeuLeuAlaAlaLeuLeuAlaLeu-----ValThrAla	19	
QY 150	AATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTAATGGCTGGTCT-----	203	
DB 20	GlnGlnArgGlyGlyGlyArgSerArgGlyGlyValLysGlySerAlaTrpGlyGlyArg	39	
QY 204	---TCTGATGAAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATG	260	
DB 40	ProAlaProPheArgSerTrpAspThrAlaArgTyrArgProTrpGlnGluGlyThrAla	59	
QY 261	AGGTGGAATAAATCTCTGGAAGGAGGCGGTGTGCGAGCGGCTCCTGACCACTGACTCACA	320	
DB 60	ArgGlnAsnAspCysTrpArgGlyGlyAspValThrPheAspIleSerAsnAspAlaPro	79	
QY 321	GCCCTCGTGGGTCAATATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAG	380	
DB 80	ThrLeuValGlyAlaArgAlaThrPheSerIleAlaLeuArgPheProGlyThrGlnThr	99	
QY 381	GAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC-----AGAAATGAGGCTGGT	434	
DB 100	ValLeuProAspGlyArgValValTrpSerGlnAsnCysThrValAsnGlyThrArgMet	119	
QY 435	TTATCTGCTGATCCATATGTTTACAACATGTCAGACAGCATGTCAGAGGACAGTACGCGGAA	494	
DB 120	LeuGlnGlyAspPro---ValTyrPro---GluGlnLeuAlaGluGlySerAspGly---	136	
QY 495	AATGACACCGGCCAAAGCCATCATACGTTCTCCCTGATGGAACCTTTCTCCTCACCCAC	554	
DB 137	-----ValPheProAspGlyGlnProPheProArgSer	147	
QY 555	CCCGATGAGAGAAAGATGGAATTTCACTACGTTCTCCACACACTTGGTCAGTATTTCCAG	614	
DB 148	AlaTrpGlyLysArgGlyArgPheValTyrValTrpTrpThrTrpGlyArgTyrTrpGln	167	
QY 615	AAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGCCT	674	
DB 168	ValValAspGlyAlaThrSerGlnLeuThrValGlyThrAspGlyValAlaLeuGlySer	187	
QY 675	CAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGG---GCATATGTTCCCATCGCA	731	
DB 188	TyrThrMetGluValValValTyrHisTyrArgGlyArgGlnArgPheIleProIleGly	207	
QY 732	CAAGTGAAGATGTGACGTGGTAACAGATCAGATTCTCTGTGTTTGTGACTATGTTCCAG	791	
DB 208	HisAlaSerThrGlnPheSerIleThrAspGlnValProIleAlaValAspValThrGln	227	
QY 792	AA-----GAACGATCGAATTCATCCGACGAAACCTTCCCAAGATCTC	835	
DB 228	LeuGluValAlaAlaGlyAspGlyGlySerPheValArgAsnArg-----	242	
QY 836	CCCATTATGTTGATGTCCTGATTTCATGATCTAGCCACTTCTCCTCAATTATCTACCATT	895	
DB 243	ProValAlaPheAsnValArgLeuHisAspProSerHisTyrLeuArgAspAlaAspIle	262	
QY 896	AACTACAAGTGGAGCTTCGGGGATAATATCTGGCTGTTTGTTCACCAATCATCTACTGTG	955	
DB 263	SerTyrSerTrpAspPheGlyAspGlnSerGlyThrLeuIleSerArgSerProThrVal	282	
QY 956	AATCACACGTATGTGCTCAATGGAACCTTCAGCCCTTAACCTCAGTGTGAAAGTCGACGA	1015	
DB 283	ThrHisThrTyrLeuGlnAlaGlySerPheAlaAlaArgLeuValLeuGlnAlaAlaIle	302	
QY 1016	CCA-----GGACTTGTCCGCCACCGCCACCCACCCAGACCTTCAAAACCC	1063	
DB 303	ProLeuSerSerCysGlyThrSerAlaProProValValAspPro---ThrThrGlyPro	321	

QY 1064 ACCCTTCTTTAGGACCTGCTGGTGACAACCCC----- 1096
Db 322 ValProSerLeuGlyProThrAlaThrGlnProValGlyProThrGlySerGlyThrAla 341
QY 1096 ----- 1096
Db 342 ThrAlaProSerAsnLeuThrGlySerGlyThrAlaAlaAlaProGlyThrThrAlaAla 361
QY 1096 ----- 1096
Db 362 ProArgAlaSerGlyAlaProAlaGluProThrGlyValSerValAlaValLeuSerAsp 381
QY 1097 ---CTGGAGCTGAGTAGGATTCCTGAT----- 1120
Db 382 SerMetGluThrGluProLeuProAspProValLeuSerThrAlaValAlaAspAlaAla 401
QY 1120 ----- 1120
Db 402 AlaGlyThrAspProThrAlaAspProLeuProProThrSerValSerSerGlyGlyAsp 421
QY 1120 ----- 1120
Db 422 AlaProGlyThrValAlaProThrAlaValGluGlySerValAlaAlaGlyValGlyThr 441
QY 1120 ----- 1120
Db 442 AlaGluAspValAlaAlaAlaThrProGlyAlaThrAlaAlaAspValAlaValAspThr 461
QY 1120 ----- 1120
Db 462 AlaGlyAlaThrAspGlyAspAlaValGlyProThrAlaAlaAlaThrAlaGluSerIle 481
QY 1120 ----- 1120
Db 482 AlaAspProThrAlaGlyAlaThrAspGlyAspAlaValGlyAlaThrAlaGluSerIle 501
QY 1120 ----- 1120
Db 502 AlaAspProThrAlaGlyAlaThrAspGlyAspAlaValGlyProThrAlaAlaAlaThr 521
QY 1120 ----- 1120
Db 522 AlaGluSerIleAlaAspProThrAlaGlyAlaThrAlaValSerSerGlySerAlaThr 541
QY 1120 ----- 1120
Db 542 AlaGlyAlaThrAlaGluProLeuLeuLeuValLysArgGlnAlaProGluAlaGluPro 561
QY 1121 GAAAACTGCCAGATTAAACAGATATGCCCACTTTCAAGCCACCATCAATGTTAGAGGGA 1180
Db 562 ThrGlyCysValLeuTyArgTyGlyThrPheSerThrGluLeuAsnIleValGlnGly 581
QY 1181 ATCTTAGAGGTTAATCATCATCCAGATGACAGACGTCCTGATGCCGTGCCATGGCCTGAA 1240
Db 582 IleGluSerValAlaIleValGln-----ValValProAlaAlaProGluGlySer 598
QY 1241 AGCTCCCTAATAGACTTTGTCTGACCTGCCAAGGGAGCATTCACACGGAGGTCTGTACC 1300
Db 599 GlyAsnSerValGluLeuThrValThrCysGluGlySerLeuProGluGluValCysThr 618
QY 1301 ATCATTTCTGACCCCACTCGGAGATCACCCAGAACACAGTCGTGCAGCCCTGTGGATGTG 1360
Db 619 ValValAlaAlaAspAlaGluCysArgThrAlaGlnMetGlnThrCysSerAlaValAlaPro 638
QY 1361 GATGAGATGTGTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGACGTAAGTGTGTG 1420
Db 639 AlaProGlyCysGlnLeuValLeuArgGlnAspPheAsnGlnSerGlyLeuTyCysLeu 658
QY 1421 AACCTCACCCCTGGGGATGACACAAGCCCTGGCTCTCACAGCACCCCTGATT----- 1471
Db 659 AsnValSerLeuAlaAsnGlyAsnGlyLeuAlaValAlaSerThrHisValAlaValGly 678
QY 1472 TCTGTTCTCTGACAGA 1486

Db 679 SerIleProSerArg 683
RESULT 5
PM17 HUMAN
ID PM17 HUMAN STANDARD; PRT; 661 AA.
AC P40967; Q12763; Q14448; Q14817; Q16565;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Melanocyte protein Pmel 17 precursor (Melanocyte lineage-specific antigen gp100) (Melanoma-associated ME20 antigen) (ME20M/ME20S) (ME20-M/ME20-S) (95 kDa melanocyte-specific secreted glycoprotein).
DE SILV OR PMEL17 OR D12S53E.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92021023; PubMed=1924386;
RA Kwon B.S., Chintamaneni C., Kozak C.A., Copeland N.G., Gilbert D.J., Jenkins N.A., Barton D., Francke U., Kobayashi Y., Kim K.-K.;
RT "A melanocyte-specific gene, Pmel 17, maps near the silver coat color locus on mouse chromosome 10 and is in a syntenic region on human chromosome 12.";
RT Proc. Natl. Acad. Sci. U.S.A. 88:9228-9232 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94327568; PubMed=7519602;
RA Adema G.J., de Boer A.J., Vogel A.M., Loenen W.A., Figdor C.G.;
RT "Molecular characterization of the melanocyte lineage-specific antigen gp100.";
RL J. Biol. Chem. 269:20126-20133 (1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96154052; PubMed=8592076;
RA Bailin T., Lee S.-T., Spritz R.A.;
RT "Genomic organization and sequence of D12S53E (Pmel 17), the human homologue of the mouse silver (si) locus.";
RL J. Invest. Dermatol. 106:24-27 (1996).
RN [4]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 25-53.
RX MEDLINE=94235165; PubMed=8179825;
RA Mareshe G.A., Marken J.S., Neubauer M., Aruffo A., Hellstroem I., Hellstroem K.E., Marquardt H.;
RT "Cloning and expression of the gene for the melanoma-associated ME20 antigen.";
RL DNA Cell Biol. 13:87-95 (1994).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=96314705; PubMed=8739560;
RA Kim K.K., Youn B.S., Heng H.H., Shi X.-M., Tsui L.-C., Lee Z.H., Pickard R.T., Kwon B.S.;
RT "Genomic organization and FISH mapping of human Pmel 17, the putative silver locus.";
RL Pigment Cell Res. 9:42-48 (1996).
RN [6]
RP SEQUENCE FROM N.A.
RA Vogel A.;
RL Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=2338257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences.";
CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Could be a melanogenic enzyme. Could represent an
CC oncofetal self-antigen that is normally expressed at low levels in
CC quiescent adult melanocytes but overexpressed by proliferating
CC neonatal melanocytes and during tumor growth. Release of the
CC soluble form, ME20-S, could protect tumor cells from antibody
CC mediated immunity.
CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL). THERE
CC IS ALSO A SECRETED SOLUBLE FORM, ME20-S, PROBABLY PRODUCT OF
CC PROTEOLYTIC CLEAVAGE.
CC -!- TISSUE SPECIFICITY: Preferentially expressed in melanomas. Some
CC expression was found in dysplastic nevi. Not found in normal
CC tissues nor in carcinomas.
CC -!- SIMILARITY: BELONGS TO THE PMEL-17/NMB FAMILY.
CC -!- SIMILARITY: Contains 1 PKD domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M77348; AAC60121.1; -;
DR EMBL; S73003; AAC60634.1; -;
DR EMBL; U31799; AAB00386.1; -;
DR EMBL; U31808; AAB00386.1; JOINED.
DR EMBL; U31807; AAB00386.1; JOINED.
DR EMBL; U31797; AAB00386.1; JOINED.
DR EMBL; U31798; AAB00386.1; JOINED.
DR EMBL; U01874; AAA18479.1; -;
DR EMBL; U20093; AAB19181.1; -;
DR EMBL; U19491; AAB19181.1; JOINED.
DR EMBL; M32295; AAA35930.1; ALT_INIT.
DR EMBL; BC001414; AAH01414.1; -;
DR PIR; I38400; I38400.
DR Genew; HGNC:10880; SILV.
DR MIM; 155550; -;
DR GO; GO:0005886; C:plasma membrane; TAS.
DR InterPro; IPR000601; PKD.
DR Pfam; PF00801; PKD; 1.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS50093; PKD; 1.
KW Transmembrane; Glycoprotein; Signal; Melanin biosynthesis; Repeat;
KW Antigen.
FT SIGNAL 1 24
FT CHAIN 25 661 MELANOCYTE PROTEIN PMEL 17.
FT DOMAIN 25 595 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 596 616 POTENTIAL.
FT DOMAIN 617 661 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 255 292 PKD.
FT DOMAIN 315 444 10 X 13 AA APPROXIMATE TANDEM REPEATS.
FT REPEAT 315 327 1.
FT REPEAT 328 340 2.
FT REPEAT 341 353 3.
FT REPEAT 354 366 4.
FT REPEAT 367 379 5.
FT REPEAT 380 392 6.
FT REPEAT 393 405 7.
FT REPEAT 406 418 8.

FT REPEAT 419 431 9.
FT REPEAT 432 444 10.
FT CARBOHYD 81 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 106 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 111 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 321 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 568 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 274 L -> P (IN REF. 1 AND 5).
FT CONFLICT 587 P -> PVPGILLT (IN REF. 1).
FT CONFLICT 592 G -> GG (IN REF. 4).
FT CONFLICT 597 P -> R (IN REF. 1).
FT CONFLICT 642 RIFCSCPIGENSPILLSQQV -> ASSALVPLVRIAPSSVG
FT SRSEYSHMML (IN REF. 6).
SQ SEQUENCE 661 AA; 70255 MW; 8A904FAB16715653 CRC64;

Alignment Scores:
Pred. No.: 1.94e-42 Length: 661
Score: 593.00 Matches: 167
Percent Similarity: 39.45% Conservative: 93
Best Local Similarity: 25.34% Mismatches: 199
Query Match: 12.23% Indels: 201
Db: 1 Gaps: 17

US-10-039-272-1 (1-2661) x PM17_HUMAN (1-661)

QY 66 TGTCTCTACTATTTCCTGGGATTTCCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCC 125
Db 8 CysLeu-----LeuHisLeuAlaValIleGlyAlaLeuAlaValGlyAlaThr 24

QY 126 AAA-----CGATTTTCATGATGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 25 LysValProArgAsnGlnAspTrpLeuGlyValSerArg----- 37

QY 180 CACAATCAATTAATGGCTGGTCTTCTGATGAAAATGACTGGATGAAAAAAGTCTACCCA 239
Db 38 -----GlnLeuArg-----ThrLysAlaTrpAsnArgGlnLeuTyrPro 50

QY 240 GTGTGGAAGCGGGAGACATGAGTGGGAAAAAACTCCTGGAAGGGAGCGCTGTGCAGGCG 299
Db 51 GluTrpThr-----GluAlaGlnArgLeuAspCysTrpArgGlyGlyGlnValSerLeu 68

QY 300 GTCCTGACCAGTACTCACCGCCCTCGTGGGCTCAATATATAACATTGCGGTGAACCTG 359
Db 69 LysValSerAsnAspGlyProThrLeuIleGlyAlaAsnAlaSerPheSerIleAlaLeu 88

QY 360 ATATTCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 89 AsnPheProGlySerGlnLysValLeuProAspGlyGlnValIleTrpValAsnAsnThr 108

QY 420 AGAAATGAGGCTGTTTATCTGCTGATCCCATATGTTTACAAC-----TGG--- 464
Db 109 -----IleIleAsnGlySerGlnValTrpGly 117

QY 465 -----ACAGCATGGTCAGAGGACAGTGACGGGAAAAAATGGCACCGGCCAAAGCCATCAT 518
Db 118 GlyGlnProValTyrProGlnGluThrAspAla----- 129

QY 519 AACGTCTTCCCTGATGGGAAAACCTTTTCCCTCACCCCGGATGGAGAAGATGGAATTC 578
Db 130 CysIlePheProAspGlyGlyProCysProSerGlySerTrpSerGlnLysArgSerPhe 149

QY 579 ATCTACGTCTTCCACACACTTGGTCAGTATTTCAGAAAATTTGGACGATGTTCAAGTGA 638
Db 150 ValTyrValTrpLysThrTrpGlyGlnTyrTrpGlnValLeuGlyGlyProValSerGly 169

QY 639 GTTCTGTGAACACACGCCAATGTGACACTTGGGCCTCAACTCATGGAAAGTACTGTCTAC 698
Db 170 LeuSerIleGlyThrGlyArgAlaMetLeuGlyThrHisThrMetGluValThrValTyr 189

QY 699 AGAAGACATGGA---CGGCATATGTTCCCATCGCACACAGTGAAGATGTGTACGTGTA 755
Db 190 HisArgArgGlySerArgSerTyrValProLeuAlaHisSerSerSerAlaPheThrIle 209

Db	295	euGlulysArgGlnAlaProLeuAspCysValleuTyrArgTyrGlySerPheSerLeuT	315
QY	1161	CCATCACAATTGTAGAGGGAATCTTAGAGGTTAAATCATCATCCAGATGACAGACGTCCTGA	1220
Db	315	hrLeuAspIleValSer-----IleGluSerAlaGluIleLeuG	328
QY	1221	TGCCGGTGCATGGCCTGAAAGCTCCCTAATAGACTTTGCTGACCTGCCAAGGAGCA	1280
Db	328	lnAlaValSerSerSerGluGlyAspAlaPheGluLeuThrValSerCysGlnGlyGlyL	348
QY	1281	TTCCACGAGGCTCTGTACCATCATTTCTGACCCACCTGCGAGATCACCCAGAACACAG	1340
Db	348	euProLysGluAlaCysMetAspIleSerSerProGlyCysGlnLeuProAlaGlnArgL	368
QY	1341	TCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCTGACTGTGTGAGACGAACTTC	1397
Db	368	euCysGlnProValProProSerProAlaCysGlnLeuValLeuHisGlnValLeuLysG	388
QY	1398	ATGGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGATGACACAAGCCTGGCTCTCA	1457
Db	388	lyGlySerGlyThrTyrCysLeuAsnValSerLeuAlaAspAlaAsnSerLeuAlaMetV	408
QY	1458	CGAGCACCCCTGATTTCTGTCTCTGACAGACCCAGCCCTCGCTTAAAGGATGGCAAACA	1517
Db	408	alSerThrGlnLeuValMetProGlyGlnGlu-----AlaGlyLeuArgGlnAlaPro-	425
QY	1518	GTGCCCTGATCTCCGTTGGCTGCTGGCCATATTTGTCACT---GTGATCTCCCTCTTG	1574
Db	426	-----LeuPheValGlyIleLeuValLeuThrAlaLeuLeuAlaSerLeuI	443
QY	1575	TGTACAAAAA 1585	
Db	443	leTyrArgArg 446	
RESULT 7			
PTN4	HUMAN	STANDARD;	PRT; 926 AA.
ID	PTN4_HUMAN		
AC	P29074;		
DT	01-DEC-1992 (Rel. 24, Created)		
DT	01-DEC-1992 (Rel. 24, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, last annotation update)		
DE	Protein tyrosine phosphatase, non-receptor type 4 (EC 3.1.3.48)		
DE	(Protein-tyrosine phosphatase MEG1) (PTPase-MEG1) (MEG).		
GN	PTPN4.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=91288564; PubMed=1648233;		
RA	Gu M., York J.D., Warshawsky I., Majerus P.W.;		
RT	"Identification, cloning, and expression of a cytosolic megakaryocyte		
RT	protein-tyrosine-phosphatase with sequence homology to cytoskeletal		
RL	protein 4.1.";		
RN	Proc. Natl. Acad. Sci. U.S.A. 88:5867-5871(1991).		
RP	[2]		
RC	SEQUENCE FROM N.A.		
RX	TISSUE=Colon;		
RA	MEDLINE=22388257; PubMed=12477932;		
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,		
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,		
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,		
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,		

Alignment Scores:	4.22e-29	Length:	491
Pred. No.:	436.50	Matches:	124
Score:	40.22%	Conservative:	63
Percent Similarity:	26.67%	Mismatches:	123
Best Local Similarity:	9.00%	Indels:	155
Query Match:	1	Gaps:	14
DB:			
US-10-039-272-1 (1-2661) x PM17_BOVIN (1-491)			
QY	603	CAGTATTTCCAGAAATTTGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTG	662
Db	1	GlnTyrTrpGlnValLeuGlyGlyProValSerGlyLeuSerIleGlyThrAspLysAla	20
QY	663	ACACTTGGCCCTCAACTCATGGAAGTACTGTCTACAGAAGACATGGA---CGGGCATAT	719
Db	21	MetLeuGlyThrTyrAsnMetGluValThrValTyrHisArgArgGlySerGlnSerTyr	40
QY	720	GTTCCCATCGCACAAGTGAAGATGTGTACGTGTAACAGATCAGATTCCTGTGTTGTG	779
Db	41	ValProLeuAlaHisSerSerSerAlaPheThrIleThrAspGlnValProPheSerVal	60
QY	780	ACTATGTTCCAG-----AAGAACGATCGAAATTCATCCGACGAAACC	821
Db	61	SerValSerGlnLeuGlnAlaLeuAspGlyArgAsnLysArg--PheLeuArgLysGln-	79
QY	822	TTCCCAAGATCTCCCATATGTTTGATGCTCTGATTCATGATCCTAGCCACTTCCCTCA	881
Db	80	-----ProLeuThrPheAlaLeuGlnLeuHisAspProSerGlyTyrLeuA	95
QY	882	ATTATTCTACCATTAACATAAGTGGAGCTTCGGGATAATACTGGCTGTTGTTTCCA	941
Db	95	laGlyAlaAspLeuSerTyrThrTyrAspPheGlyAspSerThrGlyThrLeuIleSerA	115
QY	942	CCAATCATACTGTGAATCACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTG	1001
Db	115	rgAlaLeuThrValThrHisThrTyrLeuGluSerGlyProValThrAlaGlnValValL	135
QY	1002	TGAAAGCTGCAGACCA-----GGACCTTGTCCGCCACCGCCACCACCACCCA	1049
Db	135	euGlnAlaAlaIleProLeuThrSerCysGlySerSerProValProGlyThrThrAspA	155
QY	1050	GA-----	1051
Db	155	rgHisValThrThrAlaGluAlaProGlyThrThrAlaGlyGlnValProThrThrGluV	175
QY	1051	-----	1051
Db	175	alMetGlyThrThrProGlyGlnValProThrAlaGluAlaProGlyThrThrValGlyT	195
QY	1051	-----	1051
Db	195	rpValProThrThrGluAspValGlyThrThrProGluGlnValAlaThrSerLysValL	215
QY	1051	-----	1051
Db	215	euSerThrThrProValGluMetProThrAlaLysAlaThrGlyArgThrProGluValS	235
QY	1052	-----CCTTCAAAACCACCCCTTCTTTAGGA-----	1078
Db	235	erThrThrGluProSerGlyThrThrValThrGlnGlyThrThrProGluLeuValGluT	255
QY	1079	-----CCTGCTGGTGACAAAC-----	1094
Db	255	hrThrAlaGlyGluValSerThrProGluProAlaGlySerAsnThrSerSerPheMetP	275
QY	1095	CCCTGGAG-----CTGAGTAGGATTCCTGATGAA-----	1123
Db	275	roThrGluGlyThrAlaGlySerLeuSerProLeuProAspAspThrAlaThrLeuValL	295
QY	1124	-----AACTGCCAGATTAACAGATATGCCACTTTCAAGCCA	1160

QY 1103 CTGAGTAGGATCTCTGATGAAACTGCCAGATTACAGATATGGC 1147
::: |||||
Db 518 ileArgMetLysProAspGluAsn-----GlyArgPheGly 529

RESULT 8
ODP2_RAT STANDARD; PRT; 555 AA.
ID ODP2_RAT
AC P08461;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Dihydrolipoamide acetyltransferase component of pyruvate dehydrogenase
DE complex (EC 2.3.1.12) (E2) (PDC-E2) (70 kDa mitochondrial autoantigen
DE of primary biliary cirrhosis) (PBC) (Fragment).
GN DLAT.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE OF 12-555 FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Heart;
RX MEDLINE=92256482; PubMed=1581353;
RA Matuda S., Nakano K., Ohta S., Shimura M., Yamanaka T., Nakagawa S.,
RA Titani K., Miyata T.;
RT "Molecular cloning of dihydrolipoamide acetyltransferase of the rat
RT pyruvate dehydrogenase complex: sequence comparison and evolutionary
RT relationship to other dihydrolipoamide acyltransferases.";
RL Biochim. Biophys. Acta 1131:114-118(1992).
RN [2]
RP SEQUENCE OF 86-523 FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=87196380; PubMed=3571977;
RA Gershwin M.E., McKay I.R., Sturgess A., Coppel R.L.;
RT "Identification and specificity of a cDNA encoding the 70 kd
RT mitochondrial antigen recognized in primary biliary cirrhosis.";
RL J. Immunol. 138:3525-3531(1987).
CC -!- FUNCTION: The pyruvate dehydrogenase complex catalyzes the overall
CC conversion of pyruvate to acetyl-CoA and CO(2). It contains
CC multiple copies of three enzymatic components: pyruvate
CC dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2) and
CC lipoamide dehydrogenase (E3).
CC -!- CATALYTIC ACTIVITY: Acetyl-CoA + dihydrolipoamide = CoA + S-
CC acetyldihydrolipoamide.
CC -!- COFACTOR: Contains 1 covalently bound lipoyl cofactor.
CC -!- SUBUNIT: 20 TO 30 ALPHA(2)-BETA(2) TETRAMERS OF E1 + 6 HOMODIMERS
CC OF E3 + 60 COPIES OF E2.
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -!- DISEASE: PRIMARY BILIARY CIRRHOSIS (PBC) IS AN AUTOIMMUNE DISEASE
CC CHARACTERIZED BY INFLAMMATORY OBLITERATION OF INTRA-HEPATIC
CC BILE DUCT, LEADING TO LIVER CELL DAMAGE AND CIRRHOSIS. THE E2
CC COMPONENT OF PYRUVATE DEHYDROGENASE COMPLEX IS THE AUTOANTIGEN FOR
CC PBC.
CC -!- SIMILARITY: Belongs to the 2-oxoacid dehydrogenase family.
CC -!- SIMILARITY: Contains 2 lipoyl-binding domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D10655; BAA01504.1; -;
CC EMBL; D00092; BAA20956.1; -;
CC EMBL; M16075; AAA41813.1; -;
CC PIR; S21766; S21766.
CC HSSP; P10515; 1FYC.
CC InterPro; IPR001078; 2Oxoacid_dh.
CC InterPro; IPR006257; AceF_long.
CC InterPro; IPR000089; Biotin_lipoyl.
CC InterPro; IPR004167; E3_binding.

InterPro; IPR003016; Lipoyl_BS.
Pfam; PF00198; 2-oxoacid dh; 1.
Pfam; PF00364; biotin lipoyl; 2.
Pfam; PF02817; e3 binding; 1.
ProDom; PD001115; 2Oxoacid_dh; 1.
TIGRFAMs; TIGR01349; PDHac_trf_mito; 1.
PROSITE; PS00189; LIPOYL; 2.
Glycolysis; Transferase; Acyltransferase; Repeat; Mitochondrion;
KW Lipoyl.
FT NON_TER 1 1 E3 AND/OR E1 COMPONENTS BINDING DOMAIN
FT DOMAIN 267 298 (POTENTIAL).
FT DOMAIN 379 555 CATALYTIC (BY SIMILARITY).
FT DOMAIN 523 534 COA-BINDING (BY SIMILARITY).
FT BINDING 172 172 LIPOYL.
FT ACT_SITE 528 528 POTENTIAL.
FT ACT_SITE 532 532 POTENTIAL.
FT ACT_SITE 86 87 DI -> GP (IN REF. 2).
FT CONFLICT 524 555 LSCDHRVVDGAVGAQWLAEFKKYLEKPVMTLL -> HSAVI
FT CONFLICT 524 555 IELWMEQLPEPSGLL (IN REF. 2).
FT SEQUENCE 555 AA; 58764 MW; 3D7A2DCF88DDB794 CRC64;
SQ
Alignment Scores:
Pred. No.: 0.102 Length: 555
Score: 116.50 Matches: 78
Percent Similarity: 34.90% Conservative: 41
Best Local Similarity: 22.87% Mismatches: 129
Query Match: 2.40% Indels: 93
DB: 1 Gaps: 15
US-10-039-272-1 (1-2661) x ODP2_RAT (1-555)
QY 957 ATCACACGTATGTGCTCAATGGAACCTTACCTCAGCCTTAACCTCAGCTGTAAGAGTCGAGCAC 1016
Db 213 ileAlaAlaPheAlaAspTyrArgProThrGluValThrSerLeu--LysProGlnAlap 232
QY 1017 CAGGACCTTGTCGCCACCG-----CCACCACCACCCAGACCTTCAAAACCCA 1064
Db 232 roProValProProValAlaAlaValProProIleProGlnProLeuAlaProT 252
QY 1065 CCCCTTCTTAGGACCTGCTGGTGAC-----AACCCCTGGAGCTGA 1106
Db 252 hrProSerAlaAlaProAlaGlyProLysGlyArgValPheValSerProLeuAlaLysL 272
QY 1107 GTAGGATTCCTGATGAAAACTGCCAGATTACAGATATGGCCACTTTCAGGCCACCATCA 1166
Db 272 ysLeuAlaAlaGluLysGlyIleAspLeuThr-----GlnValLysGlyT 287
QY 1167 CAATTGTAGAGGA--ATCTTAGAGTTTAAATCATCATCCAGATGACAGACGTCCTGATGC 1223
Db 287 hrGlyProGluGlyArgIleIleLysLysAspIle-----AspSerPheValP 303
QY 1224 CG-----GTGCCAT 1232
Db 303 roThrLysAlaAlaProAlaAlaAlaAlaAlaProProGlyProArgValAlaProT 323
QY 1233 GGCCTGAAAGCTCCCTAATAGACTTTGTGTGACCTGCCAAGGGAGCATTCGCCAGGAGG 1292
Db 323 hrProAlaGlyValPheIleAspIleProIle-----SerAsnI 336
QY 1293 TCTGTACCATCATTTCTGACCCACCTGCGAGATCACCCAGAACACAGTCTGCAGCCCT- 1351
Db 336 leArgArgValIleAlaGln-----ArgLeuMetGlnSerLysGlnThrIleProH 353
QY 1352 -----GTGGATGTGGATGAGATGTGTCTGTCTGCTGACTGTGAGACGAACTTCA 1397
Db 353 isTyrTyrLeuSerValAspValAsnMetGlyGluValLeuValArgLysGluLeuA 373
QY 1398 AT-----GGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGATGACACAA 1445
Db 373 snLysMetLeuGluGlyLysGlyLysIleSerValAsnAspPheIleIleLysAlaSerA 393
QY 1446 GCCTGGCTCTCAGAGCACCTGTATTCTTCTCTCTGACAGAGACCCAGCC-----T 1496

Db 393 laLeuAla-----CysLeuLysValProGluAlaAsnSerSerTrpMetAspTr 409

QY 1497 CGCCTTTAGATGGCAACAGTCGCCCTGATCTCCGTT----- 1534

Db 409 hrValIleArgGlnAsnHisValValAspValSerValAlaValSerThrProAlaGlyL 429

QY 1535 -----GGCTGCTTGGCCATATTTGTCACTG 1559

Db 429 euIleThrProIleValPheAsnAlaHisIleLysGlyLeuGluThrIleAlaSerAspV 449

QY 1560 TGATCTCCCTCTTGGTGTCACAAAAACACAGGAATACAAACCAATAGAAAATAGTCCTG 1619

Db 449 alValSerLeuAlaSerLysAlaArgGluGlyLysLeuGlnProHisGluPheGlnGlyG 469

QY 1620 GGAATGTGGTCAGAACG-----AAAGGCGCTGAGTGTCTTTCTCA 1658

Db 469 lyThrPheThrIleSerAsnLeuGlyMetPheGlyIleLysAsnPheSerAlaIleIleA 489

QY 1659 ACCGTGCAAAAGCCGTGTTCTTCCCGGGAACACAGGAAGGATCCGCTACTCAAAACC 1718

Db 489 snProProGlnAlaCysIleLeuAlaIleGlyAlaSerGluAspLysLeuIleProAlaA 509

QY 1719 AAGAATTTAAGGAGTTTCTTAAATTTGACCTTGTCTGAGCTCAGTCTTTTCA 1777

Db 509 spAsnGluLysGlyPheAspValAlaSerVal-----MetSerValThrLeuSerCys 526

RESULT 9

SM5B_MOUSE STANDARD; PRT; 1093 AA.

ID SM5B_MOUSE

AC Q60519;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Semaphorin 5B precursor (Semaphorin G) (Sema G).

GN SEMA5B OR SEMAG OR SEMG.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NMRI;

RX MEDLINE=96414430; PubMed=8817451;

RA Adams R.H., Betz H., Pueschel A.W.;

RT "A novel class of murine semaphorins with homology to thrombospondin

RT is differentially expressed during early embryogenesis.";

RL Mech. Dev. 57:33-45(1996).

CC -!- FUNCTION: May act as positive axonal guidance cues.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- TISSUE SPECIFICITY: In adult, only detected in brain.

CC -!- DEVELOPMENTAL STAGE: Differentially expressed in embryonic and

CC adult tissues. Its abundance decreases from E10 to birth.

CC -!- SIMILARITY: Belongs to the semaphorin family.

CC -!- SIMILARITY: Contains 1 Sema domain.

CC -!- SIMILARITY: Contains 7 TSP type-1 domains.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; X97818; CAA66398.1; -.

DR MGD; MGI:107555; Sema5b.

DR InterPro; IPR003659; Plexin-like.

DR InterPro; IPR002165; plexin_repeat.

DR InterPro; IPR001627; Sema.

DR InterPro; IPR000884; TSP1.

DR InterPro; IPR008085; TSP_1.

DR Pfam; PF01437; PSI; 1.

DR Pfam; PF01403; Sema; 1.

DR Pfam; PF00090; tsp; 1; 5.

DR PRINTS; PR01705; TSP1REPEAT.

DR SMART; SM00423; PSI; 1.

DR SMART; SM00630; Sema; 1.

DR SMART; SM00209; TSP1; 4.

DR PROSITE; PS50092; TSP1; 5.

KW Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;

KW Developmental protein; Glycoprotein.

FT SIGNAL 1 19

FT CHAIN 20 1093

FT DOMAIN 20 978

FT TRANSMEM 979 999

FT DOMAIN 1000 1093

FT DOMAIN 236 518

FT DOMAIN 551 605

FT DOMAIN 606 662

FT DOMAIN 664 713

FT DOMAIN 721 776

FT DOMAIN 795 850

FT DOMAIN 852 907

FT DOMAIN 908 952

FT CARBOHYD 59 59

FT CARBOHYD 95 95

FT CARBOHYD 157 157

FT CARBOHYD 178 178

FT CARBOHYD 287 287

FT CARBOHYD 333 333

FT CARBOHYD 378 378

FT CARBOHYD 532 532

FT CARBOHYD 539 539

FT CARBOHYD 547 547

FT CARBOHYD 602 602

FT CARBOHYD 728 728

FT CARBOHYD 944 944

SQ SEQUENCE 1093 AA; 120326 MW; 29B5C9B1E8108717 CRC64;

Alignment Scores:

Pred. No.: 0.156

Score: 115.50

Percent Similarity: 30.80%

Best Local Similarity: 20.77%

Query Match: 2.38%

DB: 1

US-10-039-272-1 (1-2661) x SM5B_MOUSE (1-1093)

QY 44 TCGTGAGAATTCAGCATGGAATGTCTCTACTATTTCCTGGGATTCTGCTCCTGGCTGC 103

Db 472 SerLeuArgIleLeuHisSerAlaArgAlaLeuPheValGlyLeuSerAsp-ArgValLe 491

QY 104 AAGATTGCCACTTGATGCCGCCAACGATTTCATGAT-----GTGCTGGGCAATGA 154

Db 491 uArgIleProLeuGluArgCysSerAlaTyrHisSerGlnGlyAlaCysLeuGlyAlaAr 511

QY 155 AAGACCTTCTGCTTACATGAGGAGACAAATCAATTAATGGCTGG----- 200

Db 511 gAspProTyrCys-----GlyTrpAspGlyLysArgG1 522

QY 201 -----TCTTCTGATGAAATGACTGGAATGAAACTTACCCC 238

Db 522 nLeuCysSerThrLeuGluAspSerSerAsnMetSerLeuTrpIleGlnAsnIleThrTh 542

QY 239 AGTGTGGAAGCGGGAGACATGAGG-----TGGAAAAACTCTCGGAA 280

Db 542 rCysProValArgAsnValThrArgAspGlyGlyPheGlyProTrp---SerProTrpLy 561

QY 281 G-----GGAGGCCGTGTGCAGCGGCTCTGACCAG 310

Db 561 sProCysGluHisLeuAspGlyAspAsnSerGlySerCysLeuCysArgAlaArgSerCy 581

QY 311 TGACTCACCAGCCCTCGTGGGCTCAATATAACATTTGCGGTGAACCTGATATTCCTAG 370

Db 581 sAspSerProArg-----ProAr 587
QY 371 ATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC---AGAAATGA 427
Db 587 gCysglyGlyLeuGluCysLeuGlyProSerIleHisIleAlaAsnCysSerArgAsnG1 607
QY 428 GGCTGGTTTATCTGCTGATCCATATGTTTAACTGGACAGCATGGTTCAGAGACAGTGA 487
Db 607 yAla-----TrpThrAlaTrpSerSerTrpAlaG1 617
QY 488 CGGGGAAATGGCACCGGC-----CAAAGCCATCATAACGTCCTT 526
Db 617 nCysSerThrSerCysGlyIleGlyPheGlnValArgGlnArgSerCysSerAsn----- 635
QY 527 CCCTGATGGAAACCTTTTCCTCACCACCCCGATGGAGAAGATGGAATTTCTATCTACGT 586
Db 636 -----ProAlaProArgHisGlyGly----- 642
QY 587 CTTCCACACACTTGGTCAGTATTTCAGAA-----ATTGGGACGATGTTTCAGTGAGATTT 642
Db 643 -----ArgIleCysValGlyLysSerArgGluGluArgPh 654
QY 643 CTGT-----GAACACAGCCAAATGTGACACT-----TGGSCCTCAACTCATGGA---AG 687
Db 654 eCysAsnGluAsnThrProCysProValProIlePheTrpAlaSerTrpGlySerTrpSe 674
QY 688 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 747
Db 674 rLysCysSerAsnAsnCysGlyGlyValGlnSerArgArgSer-----CysG1 692
QY 748 ACGTGGTAACAGATCAGATTCCCTGTGTTGTGACTAT-----GT 786
Db 692 uAsnGlyAsnSerCysProGlyCysGlyValGluPheLysThrCysAsnProGluAlaCy 712
QY 787 TCCAGAAACGATCGAAATTCATCCGACGAAACCTTCCCAAAGAT-----CTCCCCAT 840
Db 712 sProGlu-----ValArgAsnThrProTrpThrProTrpLeuProVa 727
QY 841 TATGTTTGAT-----GTCCT 855
Db 727 lAsnValThrGlnGlyAlaArgGlnGluGlnArgPheArgPheThrCysArgAlaPr 747
QY 856 GATTATGATGCTAGCCACTTCTCTCAATTATCT-----ACCATTAACTACAAGTGGAG 909
Db 747 oLeuProAspPro---HisGlyLeuGlnPheGlyLysArgArgThrGluThrArgThrCy 766
QY 910 CTTGGGGATAATACTAGCCTGTT----- 934
Db 766 sProAlaAspGlyThrGlyAlaCysAspThrAspAlaLeuValGluAspLeuLeuArgSe 786
QY 935 -----GTTTCCACCAATCATACTGTGAATCACACGATGTGCTCAATGGAACTTCAGC-- 988
Db 786 rGlySerThrSerProHisThrLeuAsnGlyGlyTrpAlaThrTrpGlyProTrpSerSe 806
QY 989 -----CTTAACCTCACTGTGAAAGCTGCAGCACCCAGGACCTTGTC 1029
Db 806 rCysSerArgAspCysGluLeuGlyPheArgValArgLysArgThr-----CysTh 823
QY 1030 GCCACCGCCACCAACCCAGACCTTCAAAACCCACCTTCTTTAGGACCTGCTGTGA 1089
Db 823 rAsnProGluPro-----ArgAsnGlyGlyLeuProCysValGlyAspAlaAlaG1 840
QY 1090 C-----AACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTAA 1137
Db 840 uTyrGlnAspCysAsnPro-----GlnAlaCysProValar 852
QY 1138 CAGATAT-----GGCCA 1149
Db 852 gGlyAlaTrpSerCysTrpThrAlaTrpSerGlnCysSerAlaSerCysGlyGlyHi 872
QY 1150 CTTTCAAGCCACCATCACAATTGTAGAGGGAATCTTAGAGGTTAATCATCATCCAGATGAC 1209
Db 872 sTyrGlnArgThrArgSerCysThrSer----- 881

QY 1210 AGACGTCCTGATGCCGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTG 1269
Db 882 -----ProAlaProSerProGlyGluAspIleCysLeuGlyLeuHisThrG1 897
QY 1270 CCAAGGGAGCATTCACACGGAGGTCTGTACCATCATTTCTGACCCCACTGCGAGATCAC 1329
Db 897 uGluAlaLeuCysSerThrGlnAlaCys-Pro----- 907
QY 1330 CCAGAACACAGTCTCAGCCCTGTGGATG---TGGATGAGATGTGTCTGCTGACTGTGAG 1386
Db 908 -----GluGlyTrpSerLeuTrpSerGluTrpGlyValCysThr-----G 921
QY 1387 ACGAACCTTCAATGGGTCTGGACGCTACTGTGTGAA-----CCTCACCTGGG--- 1434
Db 921 luAspGlyAlaGlnSerArgSerArgSer-CysGluGluLeuLeuProGlyProGlyAla 940
QY 1435 -----GGATGACACACAGCCTGGCTCTCACGAGCACCTGATTTCTGTCTGACAGAGA 1488
Db 941 CysValGly-AsnSerSerGlnSerArgProCysProTyrSerGluIleProValIleLe 960
QY 1489 CCCAGCCTCGCCTTTAAGGATGGCAACAGTGCC-----CTGAT 1527
Db 960 uProAlaSerSerValGluGluThrThrSerCysGlyGlyPheAsnLeuIleHisLeuI1 980
QY 1528 CTCGTTGGCTGCTTGGCCATATTTGTCACT---GTGATCTCCCTCTTGTGTGTAC--- 1579
Db 980 eValThrGlyValSerCysPheLeuValSerGlyLeuLeuThrLeuAlaValTyrLeuSe 1000
QY 1580 -----AAAAACACAAGGAATACAAACCAATAGAAAAATAGTCTGGGAA 1623
Db 1000 rCysGlnHisCysGlnArgGlnSerGlnGluSerThrLeuValHisProAlaThrProAs 1020
QY 1624 TGTGTCAGAAAGCAAGGCGCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTTCTTCCC 1683
Db 1020 nHisLeuHisTyrLysGlyGlyThr----- 1030
QY 1684 GGGAAACCAAGGAAAGGATCCGCTACTCAAAACCAAGAAATTTAA 1729
Db 1030 oLysAsnGluLysTyrThrProMet-----GluPheLys 1041

RESULT 10

HRG HUMAN

ID HRG HUMAN STANDARD; PRT; 525 AA.

AC P04196;

DT 20-MAR-1987 (Rel. 04, Created)

DT 20-MAR-1987 (Rel. 04, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Histidine-rich glycoprotein precursor (Histidine-proline rich

DE glycoprotein) (HPRG).

GN HRG.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RX MEDLINE=86216149; PubMed=3011081;

RA Koide T., Foster D.C., Yoshitake S., Davie E.W.;

RT "Amino acid sequence of human histidine-rich glycoprotein derived

RT from the nucleotide sequence of its cDNA.";

RL Biochemistry 25:2220-2225(1986).

RN [2]

RP SEQUENCE FROM N.A.

RA Wakabayashi S., Takahashi K., Tokunaga F., Koide T.;

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 214-247 FROM N.A.

RX MEDLINE=94245171; PubMed=818234;

RA Hennis B.C., Frants R.R., Bakker E., Vossen R.H., van der Poort E.W.,

RA Blonden L.A., Cox S., Khan P.M., Spurr N.K., Kluff C.;

RT "Evidence for the absence of intron H of the histidine-rich

glycoprotein (HRG) gene: genetic mapping and in situ localization of
RT HRG to chromosome 3q28-q29.";
RL Genomics 19:195-197(1994).
RN [4]
RP SEQUENCE OF 19-27.
RC TISSUE=Plasma;
RX MEDLINE=93092937; PubMed=1459097;
RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA Hochstrasser D.F.;
RT "Plasma protein map: an update by microsequencing.";
RL Electrophoresis 13:707-714(1992).
CC -!- FUNCTION: The physiological function is not yet known. It binds
CC heme, dyes and divalent metal ions. It can inhibit rosette
CC formation and is known to interact with heparin, thrombospondin,
CC and the lysine-binding site of plasminogen. On the basis of its
CC homology with HMW kininogen, the His-rich region of this protein
CC may mediate the contact activation phase of intrinsic blood
CC coagulation cascade.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
CC -!- DOMAIN: In addition to having a high His and Pro content, this
CC protein has many internal repeats. 12 tandem repetitions of a 5-
CC residue sequence (GHHPH, consensus) form a histidine-rich region.
CC -!- SIMILARITY: Contains 2 cystatin-like domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; M13149; AAA52694.1; -;
DR EMBL; AB005803; BAA21613.1; -;
DR EMBL; Z17218; CAA78925.1; -;
DR PIR; A01287; KGHUGH.
DR SWISS-2DPAGE; P04196; HUMAN.
DR Genew; HGNC:5181; HRG.
DR MIM; 142640; -;
DR InterPro; IPR000010; Cystatin.
DR Pfam; PF00031; cystatin; 1.
DR SMART; SM00043; CY; 2.
KW Glycoprotein; Heparin-binding; Repeat; Signal; Polymorphism.
FT SIGNAL 1 18
FT CHAIN 19 525 HISTIDINE-RICH GLYCOPROTEIN.
FT DOMAIN 19 136 CYSTATIN-LIKE 1.
FT DOMAIN 137 254 CYSTATIN-LIKE 2.
FT DOMAIN 276 321 PRO-RICH.
FT DOMAIN 350 497 PRO/HIS-RICH.
FT DISULFID 24 504 BY SIMILARITY.
FT DISULFID 78 89 BY SIMILARITY.
FT DISULFID 105 126 BY SIMILARITY.
FT DISULFID 203 417 BY SIMILARITY.
FT DISULFID 218 241 BY SIMILARITY.
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 344 344 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 345 345 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 204 204 P -> S (in dbSNP:3181917).
FT FTId=VAR 014528.
SQ SEQUENCE 525 AA; 59578 MW; A2B124D6CE93114F CRC64;

Alignment Scores:
Pred. No.: 0.295 Length: 525
Score: 111.00 Matches: 51
Percent Similarity: 29.50% Conservative: 8
Best Local Similarity: 25.50% Mismatches: 51
Query Match: 2.29% Indels: 91
DB: 1 Gaps: 13

US-10-039-272-1 (1-2661) x HRG_HUMAN (1-525)
QY 1028 CCGCCACCGCCACCA-----CCACCCAGACCTTCAAAACCC 1063
Db ||||||||||||||||
297 ProProProProAspGluArgAspHisSerHisGlyProProLeuProGlnGlyPro 316
QY 1064 ACCCTTCTTTAGGACCTGCTGGTGACACACCCCTGGAGCTGAGTAGGATTCCTGATGAA 1123
Db |||||
317 ProProLeuLeu-----PrometSerCysSer----- 325
QY 1124 AACTGCCAGATTAAACAGATATGGCCA----- 1149
Db ::|||
326 SerCysGlnHisAlaThrPheGlyThrAsnGlyAlaGlnArgHisSerHisAsnAsn 345
QY 1150 -----CTTTCAAGCCAC---CATCACAAATTGTAGAGGGAATCTTAGAGGTTAACAT 1197
Db |||||
346 SerSerAspLeuHisProHisLysHisSerHisGluGlnHisProHisGly---His 364
QY 1198 CATCCAGATGACAGACGCTCCTGATCCGGTGCCATGGCCCTGAAAGCTCCCTAATAGACTT 1257
Db |||||
365 HisProHisAlaHisHisProHisGlu----- 373
QY 1258 TGTGCTGACCTGCCAAGGAGGACATTCCACGGAGGTCTGTACCATCATTTCTGACCCAC 1317
Db |||||
374 ---HisAspThrHisArgGlnHisProHisGly-----HisHis-----ProHis 387
QY 1318 CTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCT 1377
Db |||||
388 -----GlyHisHisProHisGlyHisHisProHisGly----- 398
QY 1378 GACTGTGAGACGAACTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCTGGGGGA 1437
Db |||||
399 -----HisHisProHisGly 403
QY 1438 TGACACAAGCTGCTCTCAGAGACCCCTGATTTCTGTTCTGACAGAGACCCAGCCTC 1497
Db |||||
404 ---HisHisProHisCysHisAsp----- 410
QY 1498 GCCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTAC 1557
Db |||||
411 ---PheGlnAspTyrGlyProCysAspProProHisAsnGlnGlyHisCysCysHis 429

RESULT 11
PCLO MOUSE
ID PCLO MOUSE STANDARD; PRT; 5038 AA.
AC Q9QYX7; Q9QYX6; Q9QZJ0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Piccolo protein (Presynaptic cytomatrix protein) (Aczonin) (Brain-
GN derived HLMN protein).
OS PCLO OR ACZ.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, ALTERNATIVE SPLICING,
RP TISSUE SPECIFICITY, AND INTERACTION WITH PROFILIN.
RC TISSUE=Brain;
RX MEDLINE=99439764; PubMed=10508862;
RA Wang X., Kibschull M., Laue M.M., Lichte B., Petrasch-Parwez E.,
RA Kilimann M.W.;
RT "Aczonin, a 550-kd putative scaffolding protein of presynaptic active
RT zones, shares homology regions with rim and bassoon and binds
RT profilin.";
RL J. Cell Biol. 147:151-162(1999).
RN [2]
RP REVISIONS.
RA Kilimann M.W.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 4502-4682 FROM N.A.

RC TISSUE=Brain;
RA Huang W., Jin W., Huang C., Chen B., Zhang J., Ju G.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP INTERACTION WITH RIMS2.
RX MEDLINE=22384373; PubMed=12401793;
RA Fujimoto K., Shibasaki T., Yokoi N., Kashima Y., Matsumoto M.,
RA Sasaki T., Tajima N., Iwanaga T., Seino S.;
RT "Piccolo, a Ca2+ sensor in pancreatic beta-cells. Involvement of
RT cAMP-GEFII.Rim2.Piccolo complex in cAMP-dependent exocytosis.";
RL J. Biol. Chem. 277:50497-50502(2002).
CC -!- FUNCTION: May act as a scaffolding protein involved in the
CC organization of synaptic active zones and in synaptic vesicle
CC trafficking.
CC -!- SUBUNIT: Interacts with Rabac1/Pral, RIMS2 and profilin.
CC -!- SUBCELLULAR LOCATION: Concentrated at the presynaptic side of
CC synaptic junctions.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q9QYX7-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9QYX7-2; Sequence=VSP_003928, VSP_003929;
CC -!- TISSUE SPECIFICITY: Highly expressed in brain. Low levels found in
CC stomach. Not detected in other tissues analyzed including adrenal
CC gland, testis and pancreas.
CC -!- DOMAIN: C2 domain 1 is involved in binding calcium and
CC phospholipids. Calcium binds with low affinity but with high
CC specificity and induces a large conformational change.
CC -!- SIMILARITY: Contains 2 C2 domains.
CC -!- SIMILARITY: Contains 1 PDZ/DHR domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; Y19185; CAB60731.2; -.
CC EMBL; Y19186; CAB60732.2; -.
CC EMBL; AF181269; AAD55786.2; -.
CC HSP; P04410; 1A25.
CC MGD; MGI:1349390; Pclo.
CC GO; GO:0045202; C:synaptic junction; IDA.
CC GO; GO:0005509; F:calcium ion binding; ISS.
CC GO; GO:0005544; F:calcium-dependent phospholipid binding; ISS.
CC GO; GO:0005522; F:profilin binding; IDA.
CC GO; GO:0019933; P:cAMP-mediated signaling; IDA.
CC GO; GO:0007010; P:cytoskeleton organization and biogenesis; IDA.
CC GO; GO:0030073; P:insulin secretion; IDA.
CC GO; GO:0017157; P:regulation of exocytosis; IDA.
CC GO; GO:0016080; P:synaptic vesicle targeting; NAS.
CC InterPro; IPR000008; C2.
CC InterPro; IPR001478; PDZ.
CC Pfam; PF00168; C2; 2.
CC Pfam; PF00595; PDZ; 1.
CC Pfam; PF05715; Zf_piccolo; 2.
CC SMART; SM00239; C2; 2.
CC SMART; SM00228; PDZ; 1.
CC PROSITE; PS00499; C2_DOMAIN_1; 1.
CC PROSITE; PS50004; C2_DOMAIN_2; 2.
CC PROSITE; PS50106; PDZ; 1.
CC Calcium/phospholipid-binding; Metal-binding; Zinc; Zinc-finger;
KW Repeat; Alternative splicing.
KW DOMAIN 371 470 10 X 10 AA TANDEM APPROXIMATE REPEATS OF
FT P-A-K-P-Q-P-Q-P-X.
FT ZN_FING 502 526 C4-TYPE (POTENTIAL).
FT ZN_FING 967 990 C4-TYPE (POTENTIAL).
FT DOMAIN 2305 2329 POLY-PRO.
FT DOMAIN 4394 4488 PDZ.

FT DOMAIN 4607 4705 C2 DOMAIN 1.
FT DOMAIN 4922 5012 C2 DOMAIN 2.
FT VARSPLIC 4829 4833 TKPTN -> SKRRK (in isoform 2).
FT VARSPLIC 4834 5038 /FTid=VSP_003928.
FT VARSPLIC 4834 5038 Missing (in isoform 2).
FT VARSPLIC 4834 5038 /FTid=VSP_003929.
SQ SEQUENCE 5038 AA; 547600 MW; DADA460CF3B40888 CRC64;

Alignment Scores:
Pred. No.: 0.642 Length: 5038
Score: 111.00 Matches: 23
Percent Similarity: 60.38% Conservative: 9
Best Local Similarity: 43.40% Mismatches: 19
Query Match: 2.29% Indels: 2
DB: 1 Gaps: 1

US-10-039-272-1 (1-2661) x PCLO_MOUSE (1-5038)
QY 926 GGCCTGTTTGTTCACCAATCATCTGTGAATCAGTATGTGCTCAATGGACCTTC 985
Db 2280 GlyLeuProValSerThrHisProSerLysSerHisProPheArgSerSer----- 2297
QY 986 AGCCTTAACCTCACTGTGAAAGCTGCAGCACACGACCTTGTCCGCCACGCCACCA 1045
Db 2298 SerLeuAspIleSerAlaGlnProProProProProProProProProPro 2317
QY 1046 CCCAGACCTTCAAAACCCACCTTCTTTAGGACCTGCT 1084
Db 2318 ProProProProProProProProProProProProProProProProAla 2330

RESULT 12
POLR KYMVJ STANDARD; PRT; 1874 AA.
ID POLR KYMVJ
AC P36304;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE RNA replicase polypeptide (EC 2.7.7.48).
OS Kennedy yellow mosaic virus (strain Jervis bay) (KYMV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Tymoviridae;
OC Tymovirus.
OX NCBI_TaxID=36392;
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE=90218040; PubMed=2324710;
RA Ding S., Keese P., Gibbs A.;
RT "The nucleotide sequence of the genomic RNA of Kennedy yellow mosaic
RT tymovirus-Jervis Bay isolate: relationships with potex- and
RT carlaviruses.";
RL J. Gen. Virol. 71:925-931(1990).
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D00637; BAA00532.1; -.
CC PIR; JQ0533; JQ0533.
CC MEROPS; C21.001; -.
CC InterPro; IPR008043; Peptidase C21.
CC InterPro; IPR001788; RNA_dep_RNApol2.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR007094; RNA_pol_PSVir.
CC InterPro; IPR000606; Viral_helicase1.
CC Pfam; PF05381; Peptidase C21; 1.
CC Pfam; PF00978; RNA_dep_RNApol2; 1.
CC Pfam; PF01443; Viral_helicase1; 1.
KW Transferase; RNA-directed RNA polymerase; Polyprotein; ATP-binding.

RT plakoglobin/armadillo multigene family.";

RL Differentiation 61:293-304(1997).

RL [6]

RP SEQUENCE OF 543-656 FROM N.A.

RA Wilson R., Smith A., Elliott G., Kramer J., Latreille P., Keppler D.;

RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.

RL [7]

RP SEQUENCE OF 865-1225 FROM N.A.

RX MEDLINE=97202103; PubMed=9049630;

RA Simmons A.D., Overhauser J., Lovett M.;

RT "Isolation of cDNAs from the Cri-du-chat critical region by direct

RL screening of a chromosome 5-specific cDNA library.";

RL Genome Res. 7:118-127(1997).

RL [8]

RP SEQUENCE OF 873-1225 FROM N.A.

RC TISSUE=Brain;

RA Yu W., Sarginson J., Gibbs R.A.;

RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: May be involved in neuronal cell adhesion and tissue

CC morphogenesis and integrity by regulating adhesion molecules.

CC Induces cell motility and cell scattering in response to

CC hepatocyte growth factor treatment (in vitro).

CC -!- SUBUNIT: BINDS TO E-CADHERIN AT A JUXTAMEMBRANE SITE WITHIN THE

CC CYTOPLASMIC DOMAIN (BY SIMILARITY). BINDS TO PRESENILIN-1.

CC -!- SUBCELLULAR LOCATION: Adherens junction (By similarity).

CC -!- ALTERNATIVE PRODUCTS:

CC Event=Alternative splicing; Named isoforms=2;

CC Comment=Additional isoforms seem to exist;

CC Name=1; Synonyms=NPRApa;

CC IsoId=Q9UQB3-1; Sequence=Displayed;

CC Name=2; Synonyms=NPRApc;

CC IsoId=Q9UQB3-2; Sequence=VSP_006746;

CC -!- TISSUE SPECIFICITY: Predominantly expressed in brain.

CC -!- SIMILARITY: Contains 9 ARM repeats.

CC -!- SIMILARITY: Belongs to the beta-catenin family.

CC -!- CAUTION: REF.8 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO FRAMESHIFTS

CC IN POSITION 746 TO 913.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL; U96136; AAC63103.1; -

DR EMBL; U52828; AAB96357.1; -

DR EMBL; AB013805; BAA36163.1; -

DR EMBL; U81004; AAD00453.1; -

DR EMBL; U72665; AAB68599.1; -

DR EMBL; U52351; AAB97957.1; -

DR EMBL; AC003089; AAB83940.1; -

DR EMBL; AF035302; AAB88185.1; ALT_FRAME.

DR Genew; HGNC:2516; CTNND2.

DR MIM; 604275; -

DR GO; GO:0007158; P:neuronal cell adhesion; NAS.

DR InterPro; IPR008938; ARM.

DR InterPro; IPR000225; Armadillo.

DR Pfam; PF00514; Armadillo_seg; 5.

DR SMART; SM00185; ARM; 7.

DR PROSITE; PS50176; ARM_REPEAT; 3.

KW Cell adhesion; Cytoskeleton; Structural protein;

KW Developmental protein; Repeat; Coiled coil; Alternative splicing.

FT DOMAIN 49 84

FT COILED COIL (POTENTIAL).

FT POLY-PRO.

FT POLY-LYS.

FT DOMAIN 216 226

FT POLY-PRO.

FT POLY-LYS.

FT REPEAT 811 817

FT REPEAT 394 438

FT REPEAT 540 579

FT REPEAT 582 621

FT REPEAT 626 666

FT REPEAT 682 724

FT REPEAT 728 773

FT	REPEAT	835	875	ARM 7.
FT	REPEAT	882	921	ARM 8.
FT	REPEAT	975	1018	ARM 9.
FT	VARSPPLIC	823	880	Missing (in isoform 2).
FT				/FTId=VSP_006746.
FT	CONFLICT	426	426	V -> D (IN REF. 3, 4 AND 5).
FT	CONFLICT	443	443	V -> A (IN REF. 3, 4 AND 5).
FT	CONFLICT	725	725	P -> A (IN REF. 3, 4 AND 5).
FT	CONFLICT	894	894	R -> L (IN REF. 3, 4 AND 5).
FT	CONFLICT	909	909	A -> V (IN REF. 3, 4 AND 5).
FT	CONFLICT	980	980	A -> G (IN REF. 5).
FT	CONFLICT	1102	1102	A -> G (IN REF. 4).
SQ	SEQUENCE	1225	AA; 132664	MM; 8B676CFD5AF4E65 CRC64;

Alignment Scores:
Pred. No.: 0.48
Score: 110.00
Percent Similarity: 32.43%
Best Local Similarity: 21.53%
Query Match: 2.27%
DB: 1

US-10-039-272-1	(1-2661)	x	CTD2_HUMAN	(1-1225)
QY	520	ACGTCTTCCCTGATGGGAAACCTTTCTCTCACCAACCCCGGATGGAGAAGATGGAATTTC	579	---
Db	28	ThrSerSerLeuSerProGly-----LeuAsnThrSerAsnGlyAspGlySerGluThr	45	---
QY	580	TCTACGTCTTCCACACACTTGTGTCAGTATT-----	609	---
Db	46	GluThrThrSerAlaIleLeuAlaSerValLysGluGlnGluLeuGlnPheGluArgLeu	65	---
QY	609	-----	609	---
Db	66	ThrArgGluLeuGluAlaGluArgGlnIleValAlaSerGlnLeuGluArgCysLysLeu	85	---
QY	610	---TCCAGAAATTGGGACGATGTTCACTGAGAGTTTCTGTGAACACAGCCAATGTGACAC	666	---
Db	86	GlySer-GluThrGlySerMetSerSerMetSerSerAlaGluGln-----PheG1	103	---
QY	667	TTGGGCTCACTCATGGA---AGTGACTGTCTACAGAAGACATGGACGGGCATATGTTTC	723	---
Db	103	nTrpGlnSerGlnAspGlyGlnLysAspIleGluAspGluLeuThrThrGlyLeu-----	121	---
QY	724	CCATCGCACAAAGTCAAGATGTGTACGTGTTAAACAGATCAGATTCTCTGTGTTGTGACTA	783	---
Db	122	---GluLeuValAspSerCysIleArg---SerLeuGlnGluSerGlyIleLeuAsp--	138	---
QY	784	TGTTCCAGAGAACGATCGAAATTCATCCGACGAAACCTTCCCAAGATCTCCCATATAT	843	---
Db	139	---ProGlnAspTyrSerThrGlyGluArgProSerLeuSerGlnSerAlaLeuG1	157	---
QY	844	GTTTGATGTCCTGATTCATGATCCTAGCCACTTCTCCTCAATATTCTACCATTAACATA	903	---
Db	157	n-----LeuAsnSerLysProGlu-----GlySerPheGlnTyrPr	169	---
QY	904	GTGGAGCTTCGGGGATAATACTGGCCTGTTGTTTCCACCAATCATACTGTGAATCACAC	963	---
Db	169	oAlaSerTyrHisSerAsnGlnThrLeu-----AlaLeuGlyGluThrThrProSerG1	187	---
QY	964	GTATGTGCTCAATGGAACC-----	999	---
Db	187	nLeuProAlaArgGlyThrGlnAlaArgAlaThrGlyGlnSerPheSerGlnGlyThrTh	207	---
QY	1000	TGTGAAAGCT-----GAGCACCAGGACCTTGTCCGCCACCCGCCACCCACCCAG	1050	---
Db	207	rSerArgAlaGlyHisLeuAlaGlyProGluProAlaProProProProProAr	227	---
QY	1051	A-----	1051	---
Db	227	gGluProPheAlaProSerLeuGlySerAlaPheHisLeuProAspAlaProProAlaA1	247	---
QY	1052	-----	1052	-----CCTTCAAAACCC-----AC 1065

Db 247 aAlaAlaAlaLeuTyrSerSerSerThrLeuProAlaProProArgGlyGlySe 267
QY 1066 CCCTTCTTTAGGACCTGCTGTGACAAACCCCTGGAGCTGAGTAGG-----ATTCC 1116
Db 267 rProLeuAlaAlaProGlnGlyGlySerProThrLysLeuGlnArgGlyGlySerAlaPr 287
QY 1117 TGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCACATCACAAATTGTAGA 1176
Db 287 oGlu-----GlyAlaThrTyrAlaAlaProAr 296
QY 1177 GGAATCTTAGAGGTTAAATCATCATCCAGATGACAGACGCTCTGATGCCGGTGCCATGGCC 1236
Db 296 gGly-----SerSerProLysGlnSerProSerArgLeuAlaLysSerTyrSe 312
QY 1237 TGAAAGCTCCCTAATAGACTTTGTGACCTGCCAAGGGAGCATTCCTCCACGGAGGTCTG 1296
Db 312 rThrSerSerProIleAsnIleValSerSerAlaGlyLeuSerProIleArgVal-- 331
QY 1297 TACCATCATTTCTGACCCACCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGA 1356
Db 332 -----ThrSerProPro----- 335
QY 1357 TGTGGATGAGATGTCCTGTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTG 1416
Db 336 -----ThrValGlnSerThrIleSerSerSerProIleHisG1 348
QY 1417 TGTGAACCTCACCTGGGGGATGACACAAGCTGGCTCTCACGAGCACCTGATTTCTGT 1476
Db 348 nLeuSerSerThrIleGlyThrTyrAlaThrLeuSerProThrLysargLeuValHisAl 368
QY 1477 TCCTGCACAGA 1486
Db 368 aSerGluGln 371

RESULT 14
MUC5B_HUMAN
ID MUC5B_HUMAN STANDARD; PRT; 5703 AA.
AC Q9HC84; O00447; O00573; O14985; O15494; O95291; O95451; Q14881;
AC Q99552; Q9UE28;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Mucin 5B precursor (Mucin 5 subtype B, tracheobronchial) (High
DE molecular weight salivary mucin MGI) (Sublingual gland mucin).
GN MUC5B OR MUC5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 1-1594 FROM N.A.
RA Chen Y., Di Y.P., Wu R.;
RT "Molecular cloning of the amino-terminal and 5'-flanking region of the
RT human MUC5B mucin gene.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-1325 FROM N.A.
RX MEDLINE=99009274; PubMed=9790959;
RA Offner G.D., Nunes D.P., Keates A.C., Afdhal N.H., Troxler R.F.;
RT "The amino-terminal sequence of MUC5B contains conserved
RT multifunctional D domains: implications for tissue-specific mucin
RL Biochem. Biophys. Res. Commun. 251:350-355(1998).
RN [3]
RP SEQUENCE OF 40-1324 FROM N.A.
RX MEDLINE=99023932; PubMed=9804771;
RA Desseyn J.-L., Buisine M.P., Porchet N., Aubert J.-P., Laine A.;
RT "Genomic organization of the human mucin gene MUC5B: cDNA and genomic
RT sequences upstream of the large central exon.";
RL J. Biol. Chem. 273:30157-30164(1998).
RN [4]
RP SEQUENCE OF 1326-4895 FROM N.A.

RC TISSUE=Placenta;
RX MEDLINE=97166151; PubMed=9013550;
RA Desseyn J.-L., Guyonnet-Duport V., Porchet N., Aubert J.-P.,
RA Laine A.;
RT "Human mucin gene MUC5B, the 10.7 kb large central exon encodes
RT various alternate subdomains resulting in a super-repeat. Structural
RT evidence for a 11p15.5 gene family.";
RL J. Biol. Chem. 272:3168-3178(1997).
RN [5]
RP SEQUENCE OF 4057-4480 FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=97292540; PubMed=9147051;
RA Nielsen P.A., Bennett E.P., Wandall H.H., Therkildsen M.H.,
RA Hannibal J., Clausen H.;
RT "Identification of a major human high molecular weight salivary mucin
RT (MGI) as tracheobronchial mucin MUC5B.";
RL Glycobiology 7:413-419(1997).
RN [6]
RP SEQUENCE OF 4721-5703 FROM N.A.
RC TISSUE=Gall bladder;
RX MEDLINE=97293229; PubMed=9164870;
RA Keates A.C., Nunes D.P., Afdhal N.H., Troxler R.F., Offner G.D.;
RT "Molecular cloning of a major human gall bladder mucin: complete C-
RT terminal sequence and genomic organization of MUC5B.";
RL Biochem. J. 324:295-303(1997).
RN [7]
RP SEQUENCE OF 4809-5687 FROM N.A.
RC TISSUE=Sublingual gland;
RX MEDLINE=96125355; PubMed=8554565;
RA Troxler R.F., Offner G.D., Zhang F., Iontcheva I., Oppenheim F.G.;
RT "Molecular cloning of a novel high molecular weight mucin (MGI)
RT from human sublingual gland.";
RL Biochem. Biophys. Res. Commun. 217:1112-1119(1995).
RN [8]
RP SEQUENCE OF 4859-5703 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=97347489; PubMed=9201995;
RA Desseyn J.-L., Aubert J.-P., Porchet N., Laine A.;
RT "Genomic organization of the 3 region of the human MUC5B mucin.";
RL J. Biol. Chem. 272:16873-16883(1997).
CC -!- FUNCTION: Salivary mucin that is thought to contribute to the
CC lubricating and viscoelastic properties of whole saliva.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed mainly in bronchus glands and also
CC in submaxillary glands, endocervix, gall bladder, and pancreas.
CC -!- PTM: Highly glycosylated.
CC -!- SIMILARITY: Contains 1 TIL (Trypsin inhibitory-like) domain.
CC -!- SIMILARITY: Contains 3 VWFC domains.
CC -!- SIMILARITY: Contains 4 VWFD domains.
CC -!- SIMILARITY: Contains 1 C-terminal cystine knot-like (CTCK) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF107890; AAC33673.1; -.
CC EMBL; AF086604; AAC67545.1; -.
CC EMBL; AJ004862; CAA06167.1; -.
CC EMBL; Z72496; CAA96577.1; -.
CC EMBL; X74955; CAA52910.1; -.
CC EMBL; U63836; AAB61398.1; -.
CC EMBL; U78554; AAC51344.1; -.
CC EMBL; U78552; AAC51344.1; JOINED.
CC EMBL; U78553; AAC51344.1; JOINED.
CC EMBL; U78551; AAC51343.1; -.
CC EMBL; U95031; AAB65151.1; -.
CC EMBL; Y09788; CAA70926.1; -.
CC Genew; HGNC:7516; MUC5B.
CC MIM; 600770; -.

DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR006208; Cys_knot.
DR InterPro; IPR006207; Cys_knot_C.
DR InterPro; IPR009041; PMP_inhibitor.
DR InterPro; IPR002919; TIL_Cysrich.
DR InterPro; IPR006552; VC_out.
DR InterPro; IPR001007; VWF_C.
DR InterPro; IPR001846; VWF_D.
DR Pfam; PF00007; Cys_knot; 1.
DR Pfam; PF01826; TIL; 1.
DR Pfam; PF00093; vwc; 1.
DR Pfam; PF00094; vwd; 4.
DR SMART; SM00214; VWC; 6.
DR SMART; SM00215; VWC_out; 4.
DR SMART; SM00216; WVD; 4.
DR PROSITE; PS01185; CTCK_1; 1.
DR PROSITE; PS01225; CTCK_2; 1.
DR PROSITE; PS01208; VWFC_1; 2.
DR PROSITE; PS0184; VWFC_2; 2.
KW Glycoprotein; Repeat; Signal; Polymorphism.
FT SIGNAL 1 25
FT CHAIN 26 5703
FT DOMAIN 77 225
FT DOMAIN 329 386
FT DOMAIN 426 580
FT DOMAIN 858 918
FT DOMAIN 1044 1044
FT DOMAIN 1457 1603
FT DOMAIN 1609 4873
FT DOMAIN 5005 5178
FT DOMAIN 5353 5425
FT DOMAIN 5462 5528
FT DOMAIN 5594 5683
FT DISULFID 5594 5646
FT DISULFID 5622 5660
FT DISULFID 5626 5676
FT DISULFID 5643 5678
FT DISULFID 5645 5682
FT CARBOHYD 145 145
FT CARBOHYD 201 201
FT CARBOHYD 254 254
FT CARBOHYD 402 402
FT CARBOHYD 516 516
FT CARBOHYD 806 806
FT CARBOHYD 930 930
FT CARBOHYD 1277 1277
FT CARBOHYD 1293 1293
FT CARBOHYD 1557 1557
FT CARBOHYD 1775 1775
FT CARBOHYD 2192 2192
FT CARBOHYD 2721 2721
FT CARBOHYD 3419 3419
FT CARBOHYD 3948 3948
FT CARBOHYD 4745 4745
FT CARBOHYD 4901 4901
FT CARBOHYD 4958 4958
FT CARBOHYD 4965 4965
FT CARBOHYD 4987 4987
FT CARBOHYD 5037 5037
FT CARBOHYD 5052 5052
FT CARBOHYD 5156 5156
FT CARBOHYD 5427 5427
FT CARBOHYD 5467 5467
FT CARBOHYD 5506 5506
FT CARBOHYD 5507 5507
FT CARBOHYD 5543 5543
FT CARBOHYD 5553 5553
FT CARBOHYD 5604 5604
FT CARBOHYD 5618 5618
FT CARBOHYD 5662 5662
FT VARIANT 5137 5137
FT CONFLICT 34 34

FT CONFLICT 95 100
FT CONFLICT 104 104
FT CONFLICT 142 142
FT CONFLICT 225 225
FT CONFLICT 330 331
FT CONFLICT 337 337
FT CONFLICT 356 356
FT CONFLICT 362 362
FT CONFLICT 369 369
FT CONFLICT 374 374
FT CONFLICT 393 393
FT CONFLICT 468 468
FT CONFLICT 512 512
FT CONFLICT 585 587
FT CONFLICT 601 601
FT CONFLICT 628 629
FT CONFLICT 633 633
FT CONFLICT 676 676
FPGLCN -> LPCLCK (IN REF. 2).
S -> C (IN REF. 2).
E -> K (IN REF. 1).
R -> S (IN REF. 2).
PL -> T (IN REF. 2).
E -> N (IN REF. 2).
E -> K (IN REF. 2).
G -> R (IN REF. 2).
MISSING (IN REF. 2 AND 3).
D -> N (IN REF. 2).
RT -> TR (IN REF. 2).
RK -> GR (IN REF. 2).
L -> P (IN REF. 2).
GAA -> AH (IN REF. 3).
A -> S (IN REF. 3).
DP -> RS (IN REF. 2).
F -> L (IN REF. 2).
A -> P (IN REF. 3).
Alignment Scores:
Pred. No.: 0.901 Length: 5703
Score: 109.50 Matches: 118
Percent Similarity: 33.64% Conservative: 67
Best Local Similarity: 21.45% Mismatches: 201
Query Match: 2.26% Indels: 165
DB: 1 Gaps: 30
US-10-039-272-1 (1-2661) x MUSB_HUMAN (1-5703)
QY 307 CCAGTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTGCGGTGAACCTGATATTCC 366
Db 4693 ProileProSerSerThrLeuTrp-----ThrThrTrpThr 4704
QY 367 CTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATG 426
Db 4705 ValProAlaGlnThrThrProMetSerThrMetSerThrIleHisThrSerSerThr 4724
QY 427 AGGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAGGACAGTG 486
Db 4725 ProGluThrThrHisThrSerThrValLeuThrThrThrAlaThrMetThrArgAlaThr 4744
QY 487 ACGGGGAAAATGGCACCGGCCAAAGCCATCATAACGCTTTCCTGATGGGAAACCTTTTC 546
Db 4745 AsnSerThrAlaThrProSerSerThrLeuGlyThrThrArgIleLeuThrGluLeuThr 4764
QY 547 CTCACC--ACCCCGGATGGAGAGATGGA-----ATTTCATCTACGTCTTCC 591
Db 4765 ThrThrAlaThrThrThrAlaAlaThrGlySerThrAlaThrLeuSerSerThrProGly 4784
QY 592 ACACACTTGGTCAGTATTTCCAGAAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACA 651
Db 4785 Thr-ThrTrp---IleLeuThrGluProSerThrIleAlaThrValMetValProThrGl 4803
QY 652 CAGCCAATGTGACACTTGGGCCTCAACTCATGGAAGTACTGTCTACAGAAGACATGGAC 711
Db 4803 ySerThrAlaThrThr---SerSerThrLeuGlyThrAlaHisThrProLysValValTh 4822
QY 712 GGGCATATGTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTG 771
Db 4822 rAlaMetAlaThrMetProThrAla----- 4830
QY 772 TGTTTGTGACTATGTTCCAGAGAAGACGATCGAAATTCATCCGACGAAAC-----CT 822
Db 4831 -ThrAlaSerThrValProSerSerSerThrValGlyThrThrArgThrProAlaValLe 4850
QY 823 TCCCAAAGATCTCCCC-----ATTATGTTTGTATGTCTCTGATT----- 859
Db 4850 uProSerSerLeuProThrPheSerValSerThrValSerSerSerValLeuThrThrLe 4870
QY 860 -----CATGATCTAGCCACTTCCTCAATTATTCTACCAATTAACTACAAGTG 906
Db 4870 uArgProThrGlyPheProSerSerHisPhe-----SerThrProCysPheCysAr 4887

QY 907 GAGCTTCGGGAT-----AATACTGGCCTGTTTGTTCACCAATCATACTGT 954
Db :|||||
4887 gAlaPheGlyGlnPheSerProGlyGluValIleTyrAsnLysThrAspArgAlaG1 4907
QY 955 GAATCACACGTAT---GTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGC 1011
Db :|||||
4907 yCysHisPheTyrAlaValCysAsnGlnHisCysAspIleAsp-----Ar 4922
QY 1012 AGCACCAGGACCTTGTCCGCCACCGCCACCACCAGCCTTCAAAACCC-----AC 1065
Db :|||||
4922 gPheGlnGlyAlaCysProThrSerProProVal---SerSerAlaProLeuSerSe 4941
QY 1066 CCCTTCTTTAGGACCTGTGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGAT----- 1120
Db :|||||
4941 rProSerProAlaProGlyCysAspAsnAlaIleProLeuArgGlnValAsnGluThrTr 4961
QY 1121 -----GAAACTGCCAGATTAAACAGATATGGCCACTTCAAGCCACCATCAACAATTGT 1173
Db :|||||
4961 pThrLeuGluAsnCysThrValAlaArg----- 4970
QY 1174 AGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGCTGCCATG 1233
Db :|||||
4971 -----CysValGlyAspAsnArgValValLeuLeuLeuAspProLy 4983
QY 1234 GCCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGCCAAGGAGC---ATTCCCACGGA 1290
Db :|||||
4983 sProValAlaAsn-----ValThrCysValAsnLysHisLeuProIleLy 4998
QY 1291 GGTCTGTACCATCATTTCTGACCCACC-----TGCAGATCACCCAGAACACAGTCTG 1344
Db :|||||
4998 sVal-----SerAspProSerGlnProCysAspPheHisTyrGluCysGluCy 5014
QY 1345 CAGCCCTGTGGATGTGGATGAGATGTGTCTGCTG----- 1378
Db :|||||
5014 s-----IleCysSerMetTrpGlyGlySerHisTyrSerThrPh 5027
QY 1379 ----ACTGTGACGACGAACCTTCAATGGTCTGGGACGTACTGTGTG----- 1420
Db :|||||
5027 eAspGlyThrSerTyrThrPheArgGlyAsnCysThrTyrValLeuMetArgGluIleHi 5047
QY 1421 -----AACCTCACCTGGGGGATGACACA----- 1444
Db :|||||
5047 sAlaArgPheGlyAsnLeuSerLeuTyrLeuAspAsnHisTyrCysThrAlaSerAlaTh 5067
QY 1445 ----- 1452
Db :|||||
5067 rAlaAlaAlaArgCysProArgAlaLeuSerIleHisTyrLysSerMetAspIleVa 5087
QY 1453 TCTCAGGACGACCCCTGATTCT-----GTTCTGACAGAGACCCC 1491
Db :|||||
5087 lLeuThrValThrMetValHisGlyLysGluGluGlyLeuIleLeuPheAspGlnIlePr 5107
QY 1492 AGCCTGCCTTTAAGGATGGCAACACAGTCCCTGATCTCCGTT---GGCTGCTTGCCAT 1548
Db :|||||
5107 oValSerSerGlyPheSerLysAsnGlyValLeuValSerValLeuGlyThrThrMe 5127
QY 1549 ATTTGTCACGTGATCTCCCTC-----TTGGTGACAAAAA 1584
Db :|||||
5127 tArgValAspIleProAlaLeuGlyValThrValThrPheAsnGlyGlnValPheGlnAl 5147
QY 1585 ACACAAGGAATACACCAATAAGAAATAGTCCTGG---GAATGTGGT-----CA 1631
Db :|||||
5147 aArgLeuProTyrSerLeuPheHisAsnAsnThrGluGlyGlnCysGlyThrCysThrAs 5167
QY 1632 GAAGCAAAGGCTGAGTGTCTTTCTCAA 1659
Db :|||||
5167 nAsnGlnArgAspCysLeuGlnArg 5176

RESULT 15

IRF3_HUMAN

ID IRF3_HUMAN STANDARD; PRT; 427 AA.

AC Q14653;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Interferon regulatory factor 3 (IRF-3).
GN IRF3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Retina;
RX MEDLINE=96102173; PubMed=8524823;
RA Au W.W.-C., Moore P.P.A., Lowther W.W., Juang Y.-T., Pitha P.M.;
RT "Identification of a member of the interferon regulatory factor
RT family that binds to the interferon-stimulated response element and
RT activates expression of interferon-induced genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:11657-11661(1995).
RN [2]
RP SEQUENCE OF 323-413 FROM N.A.
RX MEDLINE=99020108; PubMed=9803267;
RA Bellingham J., Gregory-Evans K., Gregory-Evans C.Y.;
RT "Mapping of human interferon regulatory factor 3 (IRF3) to chromosome
RT 19q13.3-13.4 by an intragenic polymorphic marker.";
RL Ann. Hum. Genet. 62:231-234(1998).
CC -!- FUNCTION: MAY ACTIVATE TRANSCRIPTION BY COMPLEX FORMATION WITH
CC OTHER TRANSCRIPTIONAL FACTORS, POSSIBLY MEMBERS OF THE STAT
CC FAMILY. BINDS SPECIFICALLY TO THE IFN-STIMULATED RESPONSE ELEMENT
CC (ISRE) BUT NOT TO THE IRF-1 BINDING SITE PRD-I.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Expressed constitutively in a variety of
CC tissues.
CC -!- SIMILARITY: Belongs to the IRF family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z56281; CAA91227.1; -.
DR EMBL; U86636; AAC68818.1; -.
DR HSSP; P23906; 2IRF.
DR TRANSFAC; T04673; -.
DR Genew; HGNC:6118; IRF3.
DR MIM; 603734; -.
DR GO; GO:0003702; F:RNA polymerase II transcription factor acti. . .; TAS.
DR GO; GO:0003712; F:transcription cofactor activity; TAS.
DR GO; GO:0006366; P:transcription from Pol II promoter; TAS.
DR InterPro; IPR001346; IRF.
DR InterPro; IPR008984; SMAD_FHA.
DR Pfam; PF00605; IRF; 1.
DR PRINTS; PR00267; INTERNREGFCT.
DR ProDom; PD002355; IRF; 1.
DR SMART; SM00348; IRF; 1.
DR PROSITE; PS00601; IRF; 1.
DR Transcription regulation; DNA-binding; Nuclear protein; Activator;
DR Polymorphism.
FT DNA BIND 7 107 TRYPTOPHAN PENTAD REPEAT.
FT VARIANT 96 96 R -> Q (in dbSNP:968457).
FT /FTID=VAR_011901.
FT VARIANT 377 377 E -> K (in dbSNP:1049486).
FT /FTID=VAR_011902.
FT VARIANT 427 427 S -> T (in dbSNP:7251).
FT /FTID=VAR_011903.
SQ SEQUENCE 427 AA; 47219 MW; F536676FA78B0110 CRC64;

Alignment Scores:
Pred. No.: 0.665 Length: 427
Score: 106.50 Matches: 79
Percent Similarity: 33.33% Conservative: 32
Best Local Similarity: 23.72% Mismatches: 90

Search completed: September 10, 2004, 15:41:24
Job time : 87.5 secs

Query Match:	2.20%	Indels:	133
DB:	1	Gaps:	18
US-10-039-272-1 (1-2661) x IRF3_HUMAN (1-427)			
QY	383	AGATGCCAATGGCAACATAGTCTATGAGAGAACTGCAGAAATGAG---GCTGGTTTATC	439
Db	35	ArgIleProTrpLysHisGlyLeuArgGlnAspAlaGlnGln-GluAspPheGlyIle--	53
QY	440	TGCTGATCCATATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGACGGGAAATGG	499
Db	54	-----PheGlnAlaTrpAlaGlu-----A1	60
QY	500	CACCGGCCAAAGCCATCATAACTCTTCCCTGATGGGAAACCTTTTCCTCACCACCCCGG	559
Db	60	aThrGly-----AlaTyrValProGlyArgAspLysProAspLeuProTh	75
QY	560	ATGGAGAAGATGGAATTTCAT-----CTACGTCCTCCACA	594
Db	75	rTrpLysArg---AsnPheArgSerAlaLeuAsnArgLysGluGlyLeuArgLeuAlaG1	94
QY	595	CACCTGGTCAGTATTTCCTCA-----GAAATTGGGACGATGTTTCAGTGACAGATT	642
Db	94	uAspArgSerLysAspProHisAspProHisLysIleTyrGluPheValAsnSerGlyVa	114
QY	643	CTGTGAACACAGCCCAATGTGACACTTGGGCCTCAACTCATGGAAGTGACTGTCTACAGAA	702
Db	114	lGlyAspPheSerGlnProAspThrSerProAspThrAsnGlyGly-----	129
QY	703	GACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAAGATGTGTACGTGGTAACAGATC	762
Db	130	-----GlySerThrSe	133
QY	763	AGATTCCTGTGTTTGTGACTATGTTCCAGAAGAACGATCGAATTCATCCGACGAAACCT	822
Db	133	rAspThrGlnGlu--AspIleLeuAspGluLeuLeuGlyAsnMetValLeuAlaProLe	152
QY	823	TCCCAAAGATCTCCCATTTATGTTTGATGTCCTGATTTCATGATCCTAGCCACTTCCTCAA	882
Db	152	uPro-----AspPro-----	155
QY	883	TTATTCTACCATTAACATAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCAC	942
Db	155	-----	155
QY	943	CAATCATCTGTGAATCACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGT	1002
Db	156	-----GlyProSerLeu-----	160
QY	1003	GAAAGCTGCAGCACGACGACCTTGTCCGCCACCGCCACCACCACCCAGACCTTCAAAACC	1062
Db	161	-----AlaValAlaProGluProCysProGlnProLeuArgSerProSerLeuAspAsnPr	179
QY	1063	CACC-----CCTTCTTTAGGACCTGTGGTGACAAACCCCTGGAGCTGAGTAGGATTCC	1116
Db	179	oThrProPheProAsnLeuGlyProSer---GluAsnProLeuLysArgLeuLeuValPr	198
QY	1117	TGATGAAAACCTGCCAGATTAAACAGATATGGCCAC-----TTTCAAGC	1158
Db	198	oGlyGluGluTrpGluPheGluValThrAlaPheTyrArgGlyArgGlnValPheGlnG1	218
QY	1159	CACCATCACAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCT	1218
Db	218	nThiLeSerCysProGluGlyLeuArgLeuValGly---SerGluValGlyAspArgTh	237
QY	1219	GATGCCG-----GTGCCATGGCCTGAAAGCTCCCTAATAGAC-----	1255
Db	237	rLeuProGlyTrpProValThrLeuProAspProGlyMetSerLeuThrAspArgGlyVa	257
QY	1256	-----TTTGTCTGACCTGCCAAGG	1276
Db	257	lMetSerTyrValArgHisValLeuSerCysLeuGly	269

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 9, 2004, 20:27:32 ; Search time 16423 Seconds
(without alignments)
7022.822 Million cell updates/sec

Title: US-10-039-272-1
Perfect score: 2661
Sequence: 1 cggcacgagggccagagga.....aaaaaaaaaaaaaaaaaaaa 2661

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2650	99.6	2662	9	AF3222909	AF3222909 Homo sapi
2	2547.8	95.7	2669	6	AR018808	AR018808 Sequence
3	2547.8	95.7	2669	6	AR018814	AR018814 Sequence
4	2547.8	95.7	2669	6	AX818137	AX818137 Sequence
5	2547.8	95.7	2669	9	HSNMB	X76534 H.sapiens N
6	2522	94.8	2787	9	BC032783	BC032783 Homo sapi
7	2485	93.4	2683	6	AX358788	AX358788 Sequence
8	2485	93.4	2683	6	AX362281	AX362281 Sequence
9	2485	93.4	2683	9	AY359124	AY359124 Homo sapi
10	1671	62.8	1683	6	AX677738	AX677738 Sequence
11	1051.2	39.5	2282	10	BC061725	BC061725 Rattus no
12	1051.2	39.5	2303	6	BD062749	BD062749 Modulator
13	1049.8	39.5	2279	10	AF322054	AF322054 Mus muscu
14	1049.8	39.5	2299	10	BC026375	BC026375 Mus muscu
15	1047.8	39.4	2213	6	AR156839	AR156839 Sequence
16	1047.8	39.4	2213	6	BD269857	BD269857 The poly
17	1046.4	39.3	2320	10	AF184983	AF184983 Rattus no
18	1045	39.3	2305	10	MMU251685	AJ251685 Mus muscu
19	1002.6	37.7	169739	9	AC005082	AC005082 Homo sapi
20	988.2	37.1	221255	9	AC145883	AC145883 Pan trogl
21	854.2	32.1	898	11	G26743	G26743 human STS S
22	584.6	22.0	2467	5	CJQNR71	X94144 C.japonica
23	572	21.5	1690	9	BC011595	BC011595 Homo sapi
24	568.8	21.4	1593	9	HSA505015	AJ505015 Homo sapi
25	546.4	20.5	621	9	BT007074	BT007074 Homo sapi
26	546.4	20.5	621	12	BT007499	BT007499 Synthetic
27	545.6	20.5	619	6	AR176427	AR176427 Sequence
28	545.6	20.5	619	6	BD275711	BD275711 COMPOUNDS
29	545.6	20.5	619	6	BD275724	BD275724 COMPOUNDS
30	545.6	20.5	619	6	AR220496	AR220496 Sequence
31	545.6	20.5	619	6	AR220509	AR220509 Sequence
32	545.6	20.5	619	6	AR255490	AR255490 Sequence
33	545.6	20.5	619	6	AR255503	AR255503 Sequence
34	545.6	20.5	619	6	AR281060	AR281060 Sequence
35	545.6	20.5	619	6	AR281073	AR281073 Sequence
36	545.6	20.5	619	6	AR437851	AR437851 Sequence
37	545.6	20.5	619	6	AR437864	AR437864 Sequence
38	545.6	20.5	619	6	AX365712	AX365712 Sequence
39	545.6	20.5	619	6	AX365728	AX365728 Sequence
40	545.6	20.5	619	6	BD226040	BD226040 Compound
41	545.6	20.5	619	6	BD226053	BD226053 Compound
42	537.2	20.2	572	6	AX360616	AX360616 Sequence
43	526.2	19.8	570	6	AX360615	AX360615 Sequence
44	518	19.5	158105	2	AC023375	AC023375 Homo sapi
45	514.8	19.3	537	6	AX360618	AX360618 Sequence

ALIGNMENTS

RESULT 1
AF3222909
LOCUS AF3222909 2662 bp mRNA linear PRI 23-APR-2003
DEFINITION Homo sapiens transmembrane glycoprotein HGFIN mRNA, complete cds.
ACCESSION AF3222909
VERSION AF3222909.1 GI:11993664
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2662)
AUTHORS Bandari,P.S.; Qian,J.; Yehia,G.; Joshi,D.D.; Maloof,P.B.;
Potian,J.; Oh,H.S.; Gascon,P.; Harrison,J.S. and Rameshwar,P.
TITLE Hematopoietic growth factor inducible neurokinin-1 type: a

transmembrane protein that is similar to neurokinin 1 interacts
with substance P
Regul. Pept. 111 (1-3), 169-178 (2003)
MEDLINE 22498106
PUBMED 12609765
2 (bases 1 to 2662)
Rameshwar,P.
Direct Submission
Submitted (20-NOV-2000) Medicine, UMDNJ-New Jersey Medical School,
185 South Orange Ave. MSB, Rm. E-579, Newark, NJ 07103, USA
Location/Qualifiers
1. .2662
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="bone marrow peripheral blood"
60. .1742
/note="hematopoietic growth factor-inducible neurokinin-1
protein"
/codon_start=1
/product="transmembrane glycoprotein HGFIN"
/protein_id="AAG42839.1"
/db_xref="GI:1193665"
/translation="MECLYFLGLLLAARLPDAAKRFDVLGNERPSAYMREHNQL
NGWSSDENWNEKLYPVWKRGMRWKNSWKGRVQAVLTSDSPALVGSNITPAVNLIF
PRCKEDANGNIYVEKNCREAGLSADPYVYNWTAWSDESDGNGTQSHNVFPDQK
PFPHPGWRWNFIYVFTLQYFQKLGRCVRSVNTANVTLPQLMEVTVYRRHGR
AYVPIAQKDVVVTDOIPVFTFMKNDNRNSSDENFLKDLPIMFVDVLIHDSHFLNY
STINYKWSFGDNTGLFVSTNHTVNHNTYVLNGTFLSLNLTVKAAPGPCPPPPPPRPSK
PTPSLGPAGDNPLESRIPDENQCINRYHFQATITIVEGILEVNIIOQMTDVLMPVPW
PESSLIDFVVTQGSIPTEVCTIISDPTCEITQNTVCSPPVDVDEMCLLTVRRTFNGSG
TYCVNLTGLDDTSLALTSLLISVPDRDPASPLRMANSALISVGCLAI FVTVISLLVYK
KHKEYNPNIENSPGNVVRSKLSVFLNRAKAVFFPPGNOEKDPLLNQEFKGVs"

ORIGIN

Query Match 99.6%; Score 2650; DB 9; Length 2662;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2661; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 1 CGGCACGAGGCCAGAGGAATAAGTTAACCTTGGTGCCTGGCTCGGTGAGAAATTCAGCA 60
Db 1 CGGCACGAGGCCAGAGGAATAAGTTAACCTTGGTGCCTGGCTCGGTGAGAAATTCAGCA 60
QY 61 TGGAAATGCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGATG 120
Db 61 TGGAAATGCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGATG 120
QY 121 CCGCCAAACGATTCATGATGTGCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAGC 180
Db 121 CCGCCAAACGATTCATGATGTGCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAGC 180
QY 181 ACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAATACTTACCCAG 240
Db 181 ACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAATACTTACCCAG 240
QY 241 TGTGGAAGCGGGAGACATGAGGTGGAATACTCCTGGAAGGGAGGCCGTGTGCAGGCGG 300
Db 241 TGTGGAAGCGGGAGACATGAGGTGGAATACTCCTGGAAGGGAGGCCGTGTGCAGGCGG 300
QY 301 TCCTGACCAAGTACTCACCAGCCCTCGTGGCTCAAAATATAACATTTGCGGTGAACCTGA 360
Db 301 TCCTGACCAAGTACTCACCAGCCCTCGTGGCTCAAAATATAACATTTGCGGTGAACCTGA 360
QY 361 TATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGCA 420
Db 361 TATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGCA 420
QY 421 GAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAGG 480
Db 421 GAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAGG 480
QY 481 ACAGTGACGGGGAATGACCGGCCAAAGCCCATCAACCTTCCCTGATGGGAAAC 540
Db 481 ACAGTGACGGGGAATGACCGGCCAAAGCCCATCAACCTTCCCTGATGGGAAAC 540

Db 481 ACAGTGACGGGGAATGACCGGCCAAAGCCCATCAACCTTCCCTGATGGGAAAC 540
QY 541 CTTTTCCTCACCACCCCGGATGGAGAAGATGGAATTCATCTACGTCTTCCACACACTTG 600
Db 541 CTTTTCCTCACCACCCCGGATGGAGAAGATGGAATTCATCTACGTCTTCCACACACTTG 600
QY 601 GTCAGTATTTCAGAAATTTGGACGATGTTTCAGTGAGATTTCTGTGAACACAGCCAATG 660
Db 601 GTCAGTATTTCAGAAATTTGGACGATGTTTCAGTGAGATTTCTGTGAACACAGCCAATG 660
QY 661 TGACACTTGGCCCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATATG 720
Db 661 TGACACTTGGCCCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATATG 720
QY 721 TTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATTCCTGTGTTTGTGA 780
Db 721 TTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATTCCTGTGTTTGTGA 780
QY 781 CTATGTTCCAGAAACGATCGAAATTCATCCGACGAAACCTTCTCAAAGATCTCCCCA 839
Db 781 CTATGTTCCAGAAACGATCGAAATTCATCCGACGAAACCTTCTCAAAGATCTCCCCA 840
QY 840 TTATGTTTGATGTCCTGATTCATGATCCTAGCCACTTCCCTCAATATTCTACCATTAAT 899
Db 841 TTATGTTTGATGTCCTGATTCATGATCCTAGCCACTTCCCTCAATATTCTACCATTAAT 900
QY 900 ACAAGTGGAGCTTCGGGGATAATACTTGGCCTGTTGTTTCCACCAATCATACTGTGAATC 959
Db 901 ACAAGTGGAGCTTCGGGGATAATACTTGGCCTGTTGTTTCCACCAATCATACTGTGAATC 960
QY 960 ACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAAGCTGCAGCACCA 1019
Db 961 ACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAAGCTGCAGCACCA 1020
QY 1020 GACCTTGTCCGCCACCGCCACCAACCCAGACCTTCAAAAACCCACCCCTTCTTTAGGAC 1079
Db 1021 GACCTTGTCCGCCACCGCCACCAACCCAGACCTTCAAAAACCCACCCCTTCTTTAGGAC 1080
QY 1080 CTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACCTGCCAGATTAACA 1139
Db 1081 CTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACCTGCCAGATTAACA 1140
QY 1140 GATATGGCCACTTCAAGCCACCATCAAAATTTAGAGGGAATCTTAGAGGTTAAACATCA 1199
Db 1141 GATATGGCCACTTCAAGCCACCATCAAAATTTAGAGGGAATCTTAGAGGTTAAACATCA 1200
QY 1200 TCCAGATGACAGACGTCCTGATGCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTG 1259
Db 1201 TCCAGATGACAGACGTCCTGATGCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTG 1260
QY 1260 TCGTGACCTGCCAAGGGAGCATTCGCCAGGAGGTCTGTACCATCAATTTGACCCACCT 1319
Db 1261 TCGTGACCTGCCAAGGGAGCATTCGCCAGGAGGTCTGTACCATCAATTTGACCCACCT 1320
QY 1320 GCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTCTGCTGA 1379
Db 1321 GCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTCTGCTGA 1380
QY 1380 CTGTGAGACGAACCTTCAATGGGTCTGGGACGCTACTGTGTGAACCTCACCCCTGGGGATG 1439
Db 1381 CTGTGAGACGAACCTTCAATGGGTCTGGGACGCTACTGTGTGAACCTCACCCCTGGGGATG 1440
QY 1440 ACACAAGCCTGGCTCTCACGAGCACCCCTGATCTCCGTTGGCTGCTTGGCCATFATTTGTCAC 1499
Db 1441 ACACAAGCCTGGCTCTCACGAGCACCCCTGATCTCCGTTGGCTGCTTGGCCATFATTTGTCAC 1500
QY 1500 CTTTAAAGGATGGCAAAACAGTCCCTGATCTCCGTTGGCTGCTTGGCCATFATTTGTCAC 1559
Db 1501 CTTTAAAGGATGGCAAAACAGTCCCTGATCTCCGTTGGCTGCTTGGCCATFATTTGTCAC 1560
QY 1560 TGATCTCCCTCTTGGTGACAAAACACAAAGGAATACAACCCCAATAGAAAATAGTCCTG 1619
Db 1561 TGATCTCCCTCTTGGTGACAAAACACAAAGGAATACAACCCCAATAGAAAATAGTCCTG 1620

QY 1620 GGAATGTGGTCAGAAAGCAAGGCGCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCT 1679
Db |||||||
QY 1621 GGAATGTGGTCAGAAAGCAAGGCGCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCT 1680
Db |||||||
QY 1680 TCCCGGAAACCAAGGAAAGGATCCGCTACTCAAAAACCAAGAAATTAAGGAGTTTCTT 1739
Db |||||||
QY 1681 TCCCGGAAACCAAGGAAAGGATCCGCTACTCAAAAACCAAGAAATTAAGGAGTTTCTT 1740
Db |||||||
QY 1740 AAATTTCCGACCTTGTCTTCTGAAGCTCACTTTTCAGTGCCATTGATGATGATGTGCTGGA 1799
Db |||||||
QY 1741 AAATTTCCGACCTTGTCTTCTGAAGCTCACTTTTCAGTGCCATTGATGATGATGTGCTGGA 1800
Db |||||||
QY 1800 GTGGCTATTAACTTTTCTTAAGATTATTGTTAAATAGATTATTGTGTTTGGGGAA 1859
Db |||||||
QY 1801 GTGGCTATTAACTTTTCTTAAGATTATTGTTAAATAGATTATTGTGTTTGGGGAA 1860
Db |||||||
QY 1860 GTTGAATTTTATAGGTTAAATGTCACTTTTAGAGATGGGAGAGGATTATATCTGCAG 1919
Db |||||||
QY 1861 GTTGAATTTTATAGGTTAAATGTCACTTTTAGAGATGGGAGAGGATTATATCTGCAG 1920
Db |||||||
QY 1920 CAGCTTCAGCCATGTTGTGAACCTGATAAAAGCAACTTAGCAAGGCTTCTTTTCATTATT 1979
Db |||||||
QY 1921 CAGCTTCAGCCATGTTGTGAACCTGATAAAAGCAACTTAGCAAGGCTTCTTTTCATTATT 1980
Db |||||||
QY 1980 TTTTATGTTTCACTTATAAAGTCTTAGGTAAGTACTAGTAGATAAAACACCTGTGCCGAG 2039
Db |||||||
QY 1981 TTTTATGTTTCACTTATAAAGTCTTAGGTAAGTACTAGTAGATAAAACACCTGTGCCGAG 2040
Db |||||||
QY 2040 AGTAAGGAGAGAGCTACTATTGATTAGAGCCCTAACCCAGGTTAACTGCAAGAGAGCGG 2099
Db |||||||
QY 2041 AGTAAGGAGAGAGCTACTATTGATTAGAGCCCTAACCCAGGTTAACTGCAAGAGAGCGG 2100
Db |||||||
QY 2100 GGATACTTTTCAGCTTCCATGTAACTGTATGCATAAAGCCCAATGTAGTCCAGTTTCTAAG 2159
Db |||||||
QY 2101 GGATACTTTTCAGCTTCCATGTAACTGTATGCATAAAGCCCAATGTAGTCCAGTTTCTAAG 2160
Db |||||||
QY 2160 ATCATGTTCCAAAGCTAACTGAATCCCACCTTCAATACACACTCATGAACCTCTGATGGAAC 2219
Db |||||||
QY 2161 ATCATGTTCCAAAGCTAACTGAATCCCACCTTCAATACACACTCATGAACCTCTGATGGAAC 2220
Db |||||||
QY 2220 AATAACAGGCCCAAGCCTGTGGTATGATGTGACACACTTGCTAGACTCAGAAAAATACTA 2279
Db |||||||
QY 2221 AATAACAGGCCCAAGCCTGTGGTATGATGTGACACACTTGCTAGACTCAGAAAAATACTA 2280
Db |||||||
QY 2280 CTCTCATAAATGGGTGGGAGTATTTTGGTGACAACTTCTTGGCTGAGTGAAGGA 2339
Db |||||||
QY 2281 CTCTCATAAATGGGTGGGAGTATTTTGGTGACAACTTCTTGGCTGAGTGAAGGA 2340
Db |||||||
QY 2340 ATGATATTATATATTTCATTATTCATGGACATTTAGTTAGTGCTTTTATATACCAGG 2399
Db |||||||
QY 2341 ATGATATTATATATTTCATTATTCATGGACATTTAGTTAGTGCTTTTATATACCAGG 2400
Db |||||||
QY 2400 CATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTATAGTCGCTGCACATA 2459
Db |||||||
QY 2401 CATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTATAGTCGCTGCACATA 2460
Db |||||||
QY 2460 TTTGAAATCAAAATATTAAGACTTTCCAAAAATTTGGTCCCTGGTTTTTCATGGCAACTT 2519
Db |||||||
QY 2461 TTTGAAATCAAAATATTAAGACTTTCCAAAAATTTGGTCCCTGGTTTTTCATGGCAACTT 2520
Db |||||||
QY 2520 GATCAGTAAGGATTTCCCTCTGTTTGGAACTAAACCAATTTACTATATGTTAGACAAGA 2579
Db |||||||
QY 2521 GATCAGTAAGGATTTCCCTCTGTTTGGAACTAAACCAATTTACTATATGTTAGACAAGA 2580
Db |||||||
QY 2580 CATTTTCTTCTGAAAAAAAATGAGGGAAGACAAAAAATGAGGGAAGACAAAAAATGAGGGA 2639
Db |||||||
QY 2581 CATTTTCTTCTGAAAAAAAATGAGGGAAGACAAAAAATGAGGGAAGACAAAAAATGAGGGA 2640
Db |||||||
QY 2640 AAAAAAATGAGGGAAGACAAAAAATGAGGGAAGACAAAAAATGAGGGAAGACAAAAAATGAGGGA 2661
Db |||||||
QY 2641 AAAAAAATGAGGGAAGACAAAAAATGAGGGAAGACAAAAAATGAGGGAAGACAAAAAATGAGGGA 2662
Db |||||||

RESULT 2

AR018808

LOCUS

DEFINITION

AR018808

ACCESSION

AR018808

VERSION

AR018808.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches 2591; Conservative

0; Mismatches

17; Indels

3; Gaps

3;

QY 28 AACCTTGGTGCCCTGCGTGGAGAAATTCAGCATGGAATGTCTCTACTATTTCTCTGGAT 87

Db |||||||

QY 60 ACCTTGAGTGCCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCTCTGGAT 119

Db |||||||

QY 88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCGCGCCAAACGATTTTCATGATGCTGG 147

Db |||||||

QY 120 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCGCGCCAAACGATTTTCATGATGCTGG 179

Db |||||||

QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGAGCAACAATCAATTAATGGCTGCTCTCTG 207

Db |||||||

QY 180 GCAATGAAAGACCTTCTGCTTACATGAGGAGCAACAATCAATTAATGGCTGCTCTCTG 239

Db |||||||

QY 208 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 267

Db |||||||

QY 240 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 299

Db |||||||

QY 268 AAACTCCTGGAAGGAGGCGGTGTGCGAGCGGCTCTGACCACTGACTCACCAGCCCTCG 327

Db |||||||

QY 300 AAACTCCTGGAAGGAGGCGGTGTGCGAGCGGCTCTGACCACTGACTCACCAGCCCTCG 359

Db |||||||

QY 328 TGGGCTCAAAATATAACATTTGCGGTGAACCTGATATTTCCCTAGATGCCAAAGGAAGATG 387

Db |||||||

QY 360 TGGGCTCAAAATATAACATTTGCGGTGAACCTGATATTTCCCTAGATGCCAAAGGAAGATG 419

Db |||||||

QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 447

Db |||||||

QY 420 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 479

Db |||||||

QY 448 CATATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGCAGGGGAAAAATGGCACCAGCC 507

Db |||||||

QY 480 CATATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGCAGGGGAAAAATGGCACCAGCC 539

Db |||||||

QY 508 AAAGCCCATATAACGTCTTCCCTGATGGGAAACCTTTCTCACCACCCCGGATGGAGAA 567

Db |||||||

QY 540 AAAGCCCATATAACGTCTTCCCTGATGGGAAACCTTTCTCACCACCCCGGATGGAGAA 599

Db |||||||

QY 568 GATGGAATTTTCATCTACGTTCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT 627

Db |||||||

QY 600 GATGGAATTTTCATCTACGTTCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT 659

Db |||||||

QY 628 GTTCAGTGAGAGTTTCTGTGAACACAGCCAATGTGACACTTGGGCTCAACTCATGGAAG 687

Db |||||||

QY 660 GTTCAGTGAGAGTTTCTGTGAACACAGCCAATGTGACACTTGGGCTCAACTCATGGAAG 719

Db |||||||

QY 688 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 747

Db |||||||

QY 720 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 779

Db |||||||

AR018808

Sequence 91 from patent US 5783182.

AR018808

VERSION

AR018808.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

Query Match

Best Local Similarity

Matches 2591; Conservative

0; Mismatches

17; Indels

3; Gaps

3;

QY 28 AACCTTGGTGCCCTGCGTGGAGAAATTCAGCATGGAATGTCTCTACTATTTCTCTGGAT 87

Db |||||||

QY 60 ACCTTGAGTGCCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCTCTGGAT 119

Db |||||||

QY 88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCGCGCCAAACGATTTTCATGATGCTGG 147

Db |||||||

QY 120 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCGCGCCAAACGATTTTCATGATGCTGG 179

Db |||||||

QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGAGCAACAATCAATTAATGGCTGCTCTCTG 207

Db |||||||

QY 180 GCAATGAAAGACCTTCTGCTTACATGAGGAGCAACAATCAATTAATGGCTGCTCTCTG 239

Db |||||||

QY 208 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 267

Db |||||||

QY 240 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 299

Db |||||||

QY 268 AAACTCCTGGAAGGAGGCGGTGTGCGAGCGGCTCTGACCACTGACTCACCAGCCCTCG 327

Db |||||||

QY 300 AAACTCCTGGAAGGAGGCGGTGTGCGAGCGGCTCTGACCACTGACTCACCAGCCCTCG 359

Db |||||||

QY 328 TGGGCTCAAAATATAACATTTGCGGTGAACCTGATATTTCCCTAGATGCCAAAGGAAGATG 387

Db |||||||

QY 360 TGGGCTCAAAATATAACATTTGCGGTGAACCTGATATTTCCCTAGATGCCAAAGGAAGATG 419

Db |||||||

QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 447

Db |||||||

QY 420 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 479

Db |||||||

QY 448 CATATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGCAGGGGAAAAATGGCACCAGCC 507

Db |||||||

QY 480 CATATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGCAGGGGAAAAATGGCACCAGCC 539

Db |||||||

QY 508 AAAGCCCATATAACGTCTTCCCTGATGGGAAACCTTTCTCACCACCCCGGATGGAGAA 567

Db |||||||

QY 540 AAAGCCCATATAACGTCTTCCCTGATGGGAAACCTTTCTCACCACCCCGGATGGAGAA 599

Db |||||||

QY 568 GATGGAATTTTCATCTACGTTCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT 627

Db |||||||

QY 600 GATGGAATTTTCATCTACGTTCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT 659

Db |||||||

QY 628 GTTCAGTGAGAGTTTCTGTGAACACAGCCAATGTGACACTTGGGCTCAACTCATGGAAG 687

Db |||||||

QY 660 GTTCAGTGAGAGTTTCTGTGAACACAGCCAATGTGACACTTGGGCTCAACTCATGGAAG 719

Db |||||||

QY 688 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 747

Db |||||||

QY 720 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 779

Db |||||||

QY 720 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 779

Db |||||||

QY	748	ACGTGGTAACAGATCAGATTCCTGTGTGTTTGACTATGTTCCAGAAGAACGATCGAAATT	807
Db	780	ACGTGGTAACAGATCAGATTCCTGTGTGTTTGACTATGTTCCAGAAGAACGATCGAAATT	839
QY	808	CATCCGACGAAACCTTCC-CAAAGATCTCCCATTAATGTTTGATGTCCTGATTCATGATC	866
Db	840	CATCCGACGAAACCTTCCCTCAAAGATCTCCCATTAATGTTTGATGTCCTGATTCATGATC	899
QY	867	CTAGCCACTTCCTCAATTAATCTACCATTAACATAAGTGGAGCTTCGGGGATAATACTG	926
Db	900	CTAGCCACTTCCTCAATTAATCTACCATTAACATAAGTGGAGCTTCGGGGATAATACTG	959
QY	927	GCCTGTTGTTTCCACCAATCATACTGTGAATCACACGTATGTGCTCAATGGAACCTTCA	986
Db	960	GCCTGTTGTTTCCACCAATCATACTGTGAATCACACGTATGTGCTCAATGGAACCTTCA	1019
QY	987	GCCTTAACCTCACTGTGAAAGCTGCAGACACCAGGACCTTGTCCGCCACCGCCACCACCAC	1046
Db	1020	GCCTTAACCTCACTGTGAAAGCTGCAGACACCAGGACCTTGTCCGCCACCGCCACCACCAC	1079
QY	1047	CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGCTGGTGACAAACCCCTTGAGCTGA	1106
Db	1080	CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGCTGGTGACAAACCCCTTGAGCTGA	1139
QY	1107	GTAGGATTCCTGATGAAACCTGCCAGATTAACAGATATGCCCACTTTCAGGCCACCATCA	1166
Db	1140	GTAGGATTCCTGATGAAACCTGCCAGATTAACAGATATGCCCACTTTCAGGCCACCATCA	1199
QY	1167	CAATTGTAGAGGGAATCTTAGAGTTAATCATCATCCAGATGACAGACGTCTCTGATGCGGG	1226
Db	1200	CAATTGTAGAGGGAATCTTAGAGTTAATCATCATCCAGATGACAGACGTCTCTGATGCGGG	1259
QY	1227	TGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTCGTGACCTGCCAAGGGAGCAATCCCA	1286
Db	1260	TGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTCGTGACCTGCCAAGGGAGCAATCCCA	1319
QY	1287	CGGAGGTCTGTACCATCATTTCTGACCCCACTGCGAGATCAACCCAGAACACAGTCTGCA	1346
Db	1320	CGGAGGTCTGTACCATCATTTCTGACCCCACTGCGAGATCAACCCAGAACACAGTCTGCA	1379
QY	1347	GCCCTGTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGAACTTCAATGGGTCTG	1406
Db	1380	GCCCTGTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGAACTTCAATGGGTCTG	1439
QY	1407	GGACGTACTGTGTGAACCTCACCCCTGGGGATGACACAAGCCTGGCTCTCACGAGCACCC	1466
Db	1440	GGACGTACTGTGTGAACCTCACCCCTGGGGATGACACAAGCCTGGCTCTCACGAGCACCC	1499
QY	1467	TGATTTCTGTTCCCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCCTTGA	1526
Db	1500	TGATTTCTGTTCCCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCCTTGA	1559
QY	1527	TCTCCGTTGGCTGCTTGGCCATATTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAC	1586
Db	1560	TCTCCGTTGGCTGCTTGGCCATATTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAC	1619
QY	1587	ACAAGGAATACAACCCCAATAGAAAAATAGTCTCGGGAATGTGGTCAGAAGCAAGGCCTGA	1646
Db	1620	ACAAGGAATACAACCCCAATAGAAAAATAGTCTCGGGAATGTGGTCAGAAGCAAGGCCTGA	1679
QY	1647	GTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGAAAAACAGGAAAGATCCGC	1706
Db	1680	GTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGAAAAACAGGAAAGATCCGC	1739
QY	1707	TACTCAAAAAACAAGAAATTTAAAGGAGTTTCTTAAATTCGACCTTGTCTGTAAGCTCA	1766
Db	1740	TACTCAAAAAACAAGAAATTTAAAGGAGTTTCTTAAATTCGACCTTGTCTGTAAGCTCA	1799
QY	1767	CTTTTCAGTGCCATTGATGTGAGATGTGCTGGAGTGGCTATTAACTTTTTTTCCTAAAG	1826
Db	1800	CTTTTCAGTGCCATTGATGTGAGATGTGCTGGAGTGGCTATTAACTTTTTTTCCTAAAG	1859
QY	1827	ATTATTGTTAAATAGATATTGTGTTTGGGGAAGTTGAATTTTTTATAGGTTAAATGTCA	1886

Db	1860	ATTATTGTTAAATAGATATTGTGGTTTGGGGAAGTTGAATTTTTTATAGGTAAATGTCA	1919
Qy	1887	TTTTAGAGATGGGAGAGGAATTATACCTGCAGGAGCTTCAGCCATGTTGTGAAACTGAT	1946
Db	1920	TTTTAGAGATGGGAGAGGGATTATACCTGCAGGAGCTTCAGCCATGTTGTGAAACTGAT	1979
Qy	1947	AAAAGCAACTTAGCAAGGCTTCTTTTCATTATTTTTTAATGTTTCACTTATAAAGTCTTAG	2006
Db	1980	AAAAGCAACTTAGCAAGGCTTCTTTTCATTATTTTTTAATGTTTCACTTATAAAGTCTTAG	2039
Qy	2007	GTAACCTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAAAGCTACTATTGATTAA	2066
Db	2040	GTAACCTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAAAGCTACTATTGATTAA	2099
Qy	2067	GAGCCTAAAGCCAGGTTAACTGCAAGAGAGCGGGGATACCTTTCAGCTTCCATGTAACCTG	2126
Db	2100	GAGCCTAAAGCCAGGTTAACTGCAAGAGAGCGGGGATACCTTTCAGCTTCCATGTAACCTG	2159
Qy	2127	TATGCATAAAGCCAAATGTAGTCCAGTTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCA	2186
Db	2160	TATGCATAAAGCCAAATGTAGTCCAGTTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCA	2219
Qy	2187	CTTCAATACACACTCATGAACTCCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGA	2246
Db	2220	CTTCAATACACACTCATGAACTCCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGA	2279
Qy	2247	TGTGCACACTTGTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGAGTATTTTG	2306
Db	2280	TGTGCACACTTGTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGAGTATTTTG	2339
Qy	2307	GTGACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATTTCCA	2366
Db	2340	GTGACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATTTCCA	2399
Qy	2367	TGGACATTTAGTTAGTCTTTTATATATACCAGGCATGCTGAGTGACACTCTTGTGTA	2426
Db	2400	TGGACATTTAGTTAGTCTTTTATATATACCAGGCATGCTGAGTGACACTCTTGTGTA	2459
Qy	2427	TATTTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAAATATTAAGACTTTTCC	2486
Db	2460	TATTTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAAATATTAAGACTTTTCC	2518
Qy	2487	AAAAATTTGGTCCCTGGTTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTG	2546
Db	2519	AAAGATGAGGTCCTGGTTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTG	2578
Qy	2547	GAACTAAAAACCATTTACTATATGTTAGACAAGACATTTTTTTTTTCTCTCTGAAAAA	2606
Db	2579	TAACTAAAAACCATTTACTATATGTTAGACAAGACATTTTTTTTTTCTCTCTGAAAAA	2638
Qy	2607	-AAAATGAGGGAAGAGACAAAAAATAAAAAA	2636
Db	2639	TAAAGTGTGGGAAGAGACAAAAAATAAAAAA	2669

```

RESULT 3
AR018814
LOCUS          AR018814          2669 bp      DNA
DEFINITION     Sequence 101 from patent US 5783182.
ACCESSION      AR018814
VERSION        AR018814.1  GI:3973928
KEYWORDS       .
SOURCE         Unknown.
ORGANISM       Unknown.
REFERENCE      1 (bases 1 to 2669)
AUTHORS       Thompson,T.C.
TITLE         Method for identifying metastatic sequences
JOURNAL       Patent: US 5783182-A 101 21-JUL-1998;
FEATURES       Location/Qualifiers
               1..2669
               /organism="unknown"
               source

```


ORIGIN									
/mol_type="unassigned DNA"									
Query Match 95.7%; Score 2547.8; DB 6; Length 2669;									
Best Local Similarity 99.2%; Pred. No. 0;									
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;									
QY	28	AACTTGGTGCCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTACTATTTCCTGGGAT	87						
Db	60	ACCTTGAGTGCCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTACTATTTCCTGGGAT	119						
QY	88	TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGAATTCATGATGTGCTGG	147						
Db	120	TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGAATTCATGATGTGCTGG	179						
QY	148	GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGCTTCTG	207						
Db	180	GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGCTTCTG	239						
QY	208	ATGAAAAATGACTGGAATGAAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGGA	267						
Db	240	ATGAAAAATGACTGGAATGAAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGGA	299						
QY	268	AAAACTCCTGGAAGGGAGGCGGTGTCAGGCGGTCTGACAGTGACTCACCGCCCTCG	327						
Db	300	AAAACTCCTGGAAGGGAGGCGGTGTCAGGCGGTCTGACAGTGACTCACCGCCCTCG	359						
QY	328	TGGGCTCAAAATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG	387						
Db	360	TGGGCTCAAAATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG	419						
QY	388	CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC	447						
Db	420	CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC	479						
QY	448	CATATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGACGGGAAATGGCACCGGCC	507						
Db	480	CATATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGACGGGAAATGGCACCGGCC	539						
QY	508	AAAGCCATCATAACTGCTTCCCTGATGGGAAACCTTTTCTCAGAAATTCGCAAAATGGCA	567						
Db	540	AAAGCCATCATAACTGCTTCCCTGATGGGAAACCTTTTCTCAGAAATTCGCAAAATGGCA	599						
QY	568	GATGGAATTTCACTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTCGCAAAAT	627						
Db	600	GATGGAATTTCACTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTCGCAAAAT	659						
QY	628	GTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCCCTCAACTCATGGA	687						
Db	660	GTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCCCTCAACTCATGGA	719						
QY	688	TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCAAAATTCGCAAAATTCG	747						
Db	720	TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCAAAATTCGCAAAATTCG	779						
QY	748	ACGTGGTAAACAGATCAGATTCCCTGCTTTGTGACTATGTTCCAGAAATTCGCAAAATTCG	807						
Db	780	ACGTGGTAAACAGATCAGATTCCCTGCTTTGTGACTATGTTCCAGAAATTCGCAAAATTCG	839						
QY	808	CATCCGACGAAACCTTCC - CAAAGATCTCCCCATATGTTTGTGATGCTCCTGATTCATGATC	866						
Db	840	CATCCGACGAAACCTTCCCTCAAAGATCTCCCCATATGTTTGTGATGCTCCTGATTCATGATC	899						
QY	867	CTAGCCACTTCCCTCAATTAATTCACATTAACATAAGTGGAGCTTCGGGATAATACTG	926						
Db	900	CTAGCCACTTCCCTCAATTAATTCACATTAACATAAGTGGAGCTTCGGGATAATACTG	959						
QY	927	GCCTGTTGTTTCCACCAATCATACTGTGAATCACACGTATGTCTCAATGGAACTTCA	986						
Db	960	GCCTGTTGTTTCCACCAATCATACTGTGAATCACACGTATGTCTCAATGGAACTTCA	1019						
QY	987	GCCTTAACCTCACTGTGAAAGCTGCAGCACCGAGACCTTGTCCGCCACCGCCACCCAC	1046						

Db	1020	GCCTTAACCTCACTGTGAAAGCTGCAGCACCCAGGACCTTGTCCGCCACCGCCACCAC	1079
QY	1047	CCAGACCTTCAAAAACCCACCCCTTCTTTAGGACCTGCTGGTGACAAACCCCTGGAGCTGA	1106
Db	1080	CCAGACCTTCAAAAACCCACCCCTTCTTTAGGACCTGCTGGTGACAAACCCCTGGAGCTGA	1139
QY	1107	GTAGGATTCCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCA	1166
Db	1140	GTAGGATTCCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCA	1199
QY	1167	CAATTGTAGAGGGAATCTTTAGAGGTTAAACATCATCCAGATGACAGACGTCTCTGATGCCGG	1226
Db	1200	CAATTGTAGAGGGAATCTTTAGAGGTTAAACATCATCCAGATGACAGACGTCTCTGATGCCGG	1259
QY	1227	TGCCATGGCCTGAAAGCTCCCTAAATAGACTTTGTCTGACCTGCCAAGGGAGCATTTCCCA	1286
Db	1260	TGCCATGGCCTGAAAGCTCCCTAAATAGACTTTGTCTGACCTGCCAAGGGAGCATTTCCCA	1319
QY	1287	CGGAGGTCTGTACCATCATTTCTGACCCCCACCTGCGAGATCACCCAGAAACACAGTCTGCA	1346
Db	1320	CGGAGGTCTGTACCATCATTTCTGACCCCCACCTGCGAGATCACCCAGAAACACAGTCTGCA	1379
QY	1347	GCCCTGTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTG	1406
Db	1380	GCCCTGTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTG	1439
QY	1407	GGACGTACTGTGTGAACCTCACCCCTGGGGGATGACACAAGCCTGGCTCTCAGGACACCC	1466
Db	1440	GGACGTACTGTGTGAACCTCACCCCTGGGGGATGACACAAGCCTGGCTCTCAGGACACCC	1499
QY	1467	TGATTTCTGTTCCCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCCTTGA	1526
Db	1500	TGATTTCTGTTCCCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCCTTGA	1559
QY	1527	TCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGTATCTCCCTCTTGGTGTACAAAAAC	1586
Db	1560	TCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGTATCTCCCTCTTGGTGTACAAAAAC	1619
QY	1587	ACAAGGAATACAAACCCCAATAGAAAAATAGTCTCTGGGAATGTGGTCAGAAAGCAAGCCTGA	1646
Db	1620	ACAAGGAATACAAACCCCAATAGAAAAATAGTCTCTGGGAATGTGGTCAGAAAGCAAGCCTGA	1679
QY	1647	GTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCGGGAAACCCAGGAAAAGGATCCGC	1706
Db	1680	GTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCGGGAAACCCAGGAAAAGGATCCGC	1739
QY	1707	TACTCAAAAAACCAAGAAATTTAAAGAGTTTCTTAAATTTTCGACCTTGTCTTCTGAAGCTCA	1766
Db	1740	TACTCAAAAAACCAAGAAATTTAAAGAGTTTCTTAAATTTTCGACCTTGTCTTCTGAAGCTCA	1799
QY	1767	CTTTTCAGTGCCATTGATGTGAGATGTGCTGGAGTGGCTATTAACTTTTTCCTAAAG	1826
Db	1800	CTTTTCAGTGCCATTGATGTGAGATGTGCTGGAGTGGCTATTAACTTTTTCCTAAAG	1859
QY	1827	ATTATTGTAAATAGATATTGTGCTTTGGGGAAGTTGAAATTTTATAGGTTAAATGTCA	1886
Db	1860	ATTATTGTAAATAGATATTGTGCTTTGGGGAAGTTGAAATTTTATAGGTTAAATGTCA	1919
QY	1887	TTTTAGAGATGGGAGAGGGATTATACTGCAGGCAGCTTCAGCCATGTTGTGAAAACCTGAT	1946
Db	1920	TTTTAGAGATGGGAGAGGGATTATACTGCAGGCAGCTTCAGCCATGTTGTGAAAACCTGAT	1979
QY	1947	AAAAGCAACTTAGCAAGGCTCTTTTCATTATTTTTTATGTTTCACTTATAAAGCTTAG	2006
Db	1980	AAAAGCAACTTAGCAAGGCTCTTTTCATTATTTTTTATGTTTCACTTATAAAGCTTAG	2039
QY	2007	GTAACCTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAGCTACTATTGATTA	2066
Db	2040	GTAACCTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAGCTACTATTGATTA	2099
QY	2067	GAGCCTAACCCAGGTTAACTGCAAGAAGAGCGCGGATACCTTTCAGCTTTCCATGTAACCTG	2126
Db	2100	GAGCCTAACCCAGGTTAACTGCAAGAAGAGCGCGGATACCTTTCAGCTTTCCATGTAACCTG	2159

QY 268 AAAAATCCTGGAAGGGAGGCGGTGTGCAGGCGGTCTTGACCAAGTGAATCACCAGCCCTCG 327
Db |||||
290 AAAAATCCTGGAAGGGAGGCGGTGTGCAGGCGGTCTTGACCAAGTGAATCACCAGCCCTCG 349
QY 328 TGGGCTCAAAATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG 387
Db |||||
350 TGGGCTCAAAATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG 409
QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 447
Db |||||
410 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 469
QY 448 CATATGTTTACAACTGGACAGCATGGTTCAGAGCAAGTGCAGGGAAGAAATGGCACCGGCC 507
Db |||||
470 CGTATGTTTACAACTGGACAGCATGGTTCAGAGCAAGTGCAGGGAAGAAATGGCACCGGCC 529
QY 508 AAAGCCCATATAACGTCCTCCCTGATGGGAAACCTTTTCCCTCACCAACCCCGGATGGAGAA 567
Db |||||
530 AAAGCCCATATAACGTCCTCCCTGATGGGAAACCTTTTCCCTCACCAACCCCGGATGGAGAA 589
QY 568 GATGGAATTCATCTACGTCCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACCGAT 627
Db |||||
590 GATGGAATTCATCTACGTCCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACCGAT 649
QY 628 GTTCACTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG 687
Db |||||
650 GTTCACTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG 709
QY 688 TGACTGTCTACAGAAGACATGACGCGGCATATGTTCCCATCGCACAAAGTGAAGATGATG 747
Db |||||
710 TGACTGTCTACAGAAGACATGACGCGGCATATGTTCCCATCGCACAAAGTGAAGATGATG 769
QY 748 ACGTGGTAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAAGACGATCGAAAT 807
Db |||||
770 ACGTGGTAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAAGACGATCGAAAT 829
QY 808 CATCCGACGAAACCTTCC- CAAAAGATCTCCCCATTATGTTGACTATGTTCCAGAAAGACGATCGAAAT 866
Db |||||
830 CATCCGACGAAACCTTCCCAAAGATCTCCCCATTATGTTGATGTCCTGATTCATGATC 889
QY 867 CTAGCCACTTCCCTCAATTAATCTACCATTAACCTACAGTATGTCCTCAATGGAACCTTCA 926
Db |||||
890 CTAGCCACTTCCCTCAATTAATCTACCATTAACCTACAGTATGTCCTCAATGGAACCTTCA 949
QY 927 GCCTGTTGTTTCCACCAATCATACTGTGAATCACACGATGTCCTCAATGGAACCTTCA 986
Db |||||
950 GCCTGTTGTTTCCACCAATCATACTGTGAATCACACGATGTCCTCAATGGAACCTTCA 1009
QY 987 GCCTTAACCTCACTGTGAAAGCTGCAGCACAGGACCTTGTCCGCCACCGCCACCCAC 1046
Db |||||
1010 GCCTTAACCTCACTGTGAAAGCTGCAGCACAGGACCTTGTCCGCCACCGCCACCCAC 1069
QY 1047 CCAGACCTTCAAAAACCCACCCCTTCTTT----- 1074
Db |||||
1070 CCAGACCTTCAAAAACCCACCCCTTCTTTAGCAACTACTCTAAAATCTTATGATTCAAAACA 1129
QY 1075 ----AGGACCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACTGCC 1130
Db |||||
1130 CCCAGGACCTTACTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACTGCC 1189
QY 1131 AGATTAAACAGATATGGCCACTTTCAAGCCACCATCAAAATGTAGAGGGAATCTTAGAGG 1190
Db |||||
1190 AGATTAAACAGATATGGCCACTTTCAAGCCACCATCAAAATGTAGAGGGAATCTTAGAGG 1249
QY 1191 TTAACATCATCCAGATGACAGACGTCCTGATGCCGTGCCATGGCCTGAAAGCTCCCTAA 1250
Db |||||
1250 TTAACATCATCCAGATGACAGACGTCCTGATGCCGTGCCATGGCCTGAAAGCTCCCTAA 1309
QY 1251 TAGACTTTGTCGTGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACCATCATTTCTG 1310
Db |||||
1310 TAGACTTTGTCGTGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACCATCATTTCTG 1369

QY 1311 ACCCCACCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT 1370
Db |||||
1370 ACCCCACCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT 1429
QY 1371 GTCTGCTGACTGTGAGACGAACCTTCAATGGTCTGGGACGTACTGTGTGAACCTCACCC 1430
Db |||||
1430 GTCTGCTGACTGTGAGACGAACCTTCAATGGTCTGGGACGTACTGTGTGAACCTCACCC 1489
QY 1431 TGGGGGATGACACAAAGCCTGGCTCTCACGAGCACCTGATTTCTGTTCCCTGACAGAGACC 1490
Db |||||
1490 TGGGGGATGACACAAAGCCTGGCTCTCACGAGCACCTGATTTCTGTTCCCTGACAGAGACC 1549
QY 1491 CAGCCTCGCCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATAT 1550
Db |||||
1550 CAGCCTCGCCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATAT 1609
QY 1551 TTGTCACCTGATCTCCCTCTTGGTGACAAAAACACAAGGAATACAAACCAATAGAAA 1610
Db |||||
1610 TTGTCACCTGATCTCCCTCTTGGTGACAAAAACACAAGGAATACAAACCAATAGAAA 1669
QY 1611 ATAGTCTCGGGAATGTGGTCAGAAAGCAAGGCTGAGTGTCTTCTCAACCCGTGCAAAAG 1670
Db |||||
1670 ATAGTCTCGGGAATGTGGTCAGAAAGCAAGGCTGAGTGTCTTCTCAACCCGTGCAAAAG 1730
QY 1671 CCGTGTCTTCCCGGAAACCAAGGAAAGGATCCGCTACTCAAAAAACCAAGAAATTTAAAG 1730
Db |||||
1730 CCGTGTCTTCCCGGAAACCAAGGAAAGGATCCGCTACTCAAAAAACCAAGAAATTTAAAG 1789
QY 1731 GAGTTTCTTAAATTTTCGACCTGTTTCTGAAGCTCACCTTTCAGTGCCATTTGATGTGAGA 1790
Db |||||
1790 GAGTTTCTTAAATTTTCGACCTGTTTCTGAAGCTCACCTTTCAGTGCCATTTGATGTGAGA 1849
QY 1791 TGTGCTGGAGTGGCTAATAACCTTTTTCCTTAAAGATTAATTTGTTAAATAGATATTGTGG 1850
Db |||||
1850 TGTGCTGGAGTGGCTAATAACCTTTTTCCTTAAAGATTAATTTGTTAAATAGATATTGTGG 1909
QY 1851 TTTGGGGAAGTTGAATTTTATAGGTTAAATGTCAATTTTAGAGATGGGGAGAGGATTA 1910
Db |||||
1910 TTTGGGGAAGTTGAATTTTATAGGTTAAATGTCAATTTTAGAGATGGGGAGAGGATTA 1969
QY 1911 TACTGCAGGCAGCTTCAGCCATGTTGTGAACCTGATAAAAGCAACTTAGCAGATAAGAACACTG 2030
Db |||||
1970 TACTGCAGGCAGCTTCAGCCATGTTGTGAACCTGATAAAAGCAACTTAGCAGATAAGAACACTG 2089
QY 1971 TTCATTATTTTATGTTTCACTTATAAGTCTTAGGTAAGTATAGGATAAGAACACTG 2030
Db |||||
2030 TTCATTATTTTATGTTTCACTTATAAGTCTTAGGTAAGTATAGGATAAGAACACTG 2089
QY 2031 TGTCCTGAGAGTAAGGAGAGAGCTACTATTGATTAGAGCCTAACCCAGGTTAACTGCAA 2090
Db |||||
2090 TGTCCTGAGAGTAAGGAGAGAGCTACTATTGATTAGAGCCTAACCCAGGTTAACTGCAA 2149
QY 2091 GAAGCGCGGATACCTTTCAGCTTTCATGTAACCTGATGATAAGCCCAATGTAGTCCA 2150
Db |||||
2150 GAAGCGCGGATACCTTTCAGCTTTCATGTAACCTGATGATAAGCCCAATGTAGTCCA 2209
QY 2151 GTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTC 2210
Db |||||
2210 GTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTC 2269
QY 2211 TGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGTCTAGACTCAGAA 2270
Db |||||
2270 TGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGTCTAGACTCAGAA 2329
QY 2271 AAAATACTACTCTCATAAATGGGTGGGAGTATTTTGGTGACAAACCTACTTTTGGCTG 2330
Db |||||
2330 AAAATACTACTCTCATAAATGGGTGGGAGTATTTTGGTGACAAACCTACTTTTGGCTG 2389
QY 2331 AGTGAAGGAATGATATCATATATTTCATTTATTCATGGACATTTTAGTTAGCTTTTAA 2390
Db |||||
2390 AGTGAAGGAATGATATCATATATTTCATTTATTCATGGACATTTTAGTTAGCTTTTAA 2449
QY 2391 TATACCAGGCATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTATAGTCG 2450

QY 1047 CCAGACCTTCAAACCCACCCCTTCTTT----- 1074
Db 1070 CCAGACCTTCAAACCCACCCCTTCTTTAGCAACTACTCTAAATCTTATGATTCAAACA 1129
QY 1075 ----AGGACCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCC 1130
Db 1130 CCCAGGACCTACTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCC 1189
QY 1131 AGATTAAACAGATATGGCCACTTTCAAGCCACCATCAACAATGTAGAGGGAATCTTAGAGG 1190
Db 1190 AGATTAAACAGATATGGCCACTTTCAAGCCACCATCAACAATGTAGAGGGAATCTTAGAGG 1249
QY 1191 TTAACATCATCCAGATGACAGACGTCCTGATGCCGTGCCATGGCCTGAAAGCTCCCTAA 1250
Db 1250 TTAACATCATCCAGATGACAGACGTCCTGATGCCGTGCCATGGCCTGAAAGCTCCCTAA 1309
QY 1251 TAGACTTTGTGCTGACCTGCCAAGGGAGCATTCACACGGAGGTCTGTACCATCATTTCTG 1310
Db 1310 TAGACTTTGTGCTGACCTGCCAAGGGAGCATTCACACGGAGGTCTGTACCATCATTTCTG 1369
QY 1311 ACCCCACCTGCGAGATCACCCAGAACACAGTCCTGAGCCCTGTGGATGTGGATGAGATGT 1370
Db 1370 ACCCCACCTGCGAGATCACCCAGAACACAGTCCTGAGCCCTGTGGATGTGGATGAGATGT 1429
QY 1371 GTCTGCTGACTGTGAGACGAACCTTCAATGGTCTGGGACGTA CTGTGTGAAACCTCACCC 1430
Db 1430 GTCTGCTGACTGTGAGACGAACCTTCAATGGTCTGGGACGTA CTGTGTGAAACCTCACCC 1489
QY 1431 TGGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCTGTGACAGAGCC 1490
Db 1490 TGGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCTGTGACAGAGCC 1549
QY 1491 CAGCCTCGCCTTTAAGGATGGCAAACAGTCGCCCTGATCTCCGTTGGCTGTGGCCATAT 1550
Db 1550 CAGCCTCGCCTTTAAGGATGGCAAACAGTCGCCCTGATCTCCGTTGGCTGTGGCCATAT 1609
QY 1551 TTGTCACTGTGATCTCCCTCTTGGTGATACAAAAACACAAGGAATACAAACCAATAGAAA 1610
Db 1610 TTGTCACTGTGATCTCCCTCTTGGTGATACAAAAACACAAGGAATACAAACCAATAGAAA 1669
QY 1611 ATAGTCCTGGGAATGTGGTCAGAAAGCAAGCCCTGAGTGCTTTCTCAACCGTGCAAAAG 1670
Db 1670 ATAGTCCTGGGAATGTGGTCAGAAAGCAAGCCCTGAGTGCTTTCTCAACCGTGCAAAAG 1729
QY 1671 CCGTGTCTTCCCGGAAACCCAGGAAAGGATCCGCTACTCAAAAACCAAGAATTTAAAG 1730
Db 1730 CCGTGTCTTCCCGGAAACCCAGGAAAGGATCCGCTACTCAAAAACCAAGAATTTAAAG 1789
QY 1731 GAGTTCTTAAATTTTCGACCTTGTCTGAAAGCTCACTTTTCAGTGCCTATGATGTGAGA 1790
Db 1790 GAGTTCTTAAATTTTCGACCTTGTCTGAAAGCTCACTTTTCAGTGCCTATGATGTGAGA 1849
QY 1791 TGTCGGAGTGGCTATTAAACCTTTTTCCTAAAGATTTTGTAAATAGATATTGTGG 1850
Db 1850 TGTCGGAGTGGCTATTAAACCTTTTTCCTAAAGATTTTGTAAATAGATATTGTGG 1909
QY 1851 TTTGGGGAAGTTGAATTTTTTATAGGTTAAATGTCACTTTTAGAGATGGGAGAGGGATTA 1910
Db 1910 TTTGGGGAAGTTGAATTTTTTATAGGTTAAATGTCACTTTTAGAGATGGGAGAGGGATTA 1969
QY 1911 TACTGCAGGCAGCTTCAGCCATGTTGTGAAACTGTATAAAGCAACTTAGCAAGGCTTCTT 1970
Db 1970 TACTGCAGGCAGCTTCAGCCATGTTGTGAAACTGTATAAAGCAACTTAGCAAGGCTTCTT 2029
QY 1971 TTCATTATTTTTTATGTTTCACTTATAAAGTCTTAGGTAAGTGTAGGATAGAAACACTG 2030
Db 2030 TTCATTATTTTTTATGTTTCACTTATAAAGTCTTAGGTAAGTGTAGGATAGAAACACTG 2089
QY 2031 TGTCCCGAGAGTAAGGAGAGAGCTACTATTGATTAGAGCCTTAACCCAGGTTAACTGCAA 2090
Db 2090 TGTCCCGAGAGTAAGGAGAGAGCTACTATTGATTAGAGCCTTAACCCAGGTTAACTGCAA 2149

QY 2091 GAAAGAGCGGGATACTTTCAGCTTTCCATGTAACTGTATGCATAAAAGCCAATGTAGTCCA 2150
Db 2150 GAAAGAGCGGGATACTTTCAGCTTTCCATGTAACTGTATGCATAAAAGCCAATGTAGTCCA 2209
QY 2151 GTTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCACTTCAATACACACTCATGAATCC 2210
Db 2210 GTTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCACTTCAATACACACTCATGAATCC 2269
QY 2211 TGATGGAAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGTCTAGACTCAGAA 2270
Db 2270 TGATGGAAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGTCTAGACTCAGAA 2329
QY 2271 AAAATACTACTCTCATAAATGGGTGGGAGTATTTTGGTGACAACTACTTTGCTGGCTG 2330
Db 2330 AAAATACTACTCTCATAAATGGGTGGGAGTATTTTGGTGACAACTACTTTGCTGGCTG 2389
QY 2331 AGTGAAGGAATGATATTTCATATATTTCATTATTCCTATGGACATTTAGTTAGTCTTTTA 2390
Db 2390 AGTGAAGGAATGATATTTCATATATTTCATTATTCCTATGGACATTTAGTTAGTCTTTTA 2449
QY 2391 TATACCAGGCATGATGCTGAGTGACACTCTTGTGTATATTTCAAAATTTTGTATAGTCG 2450
Db 2450 TATACCAGGCATGATGCTGAGTGACACTCTTGTGTATATTTCAAAATTTTGTACAGTCG 2509
QY 2451 CTGCACATATTGAAATCAAAATATTAAAGACTTTCCAAAATTTGGTCCCTGGTTTTCA 2510
Db 2510 CTGCACATATTGAAATC-ATATATTAAAGACTTTCCAAAGATGAGGTCCCTGGTTTTCA 2568
QY 2511 TGGCAACTTGATCAGTAAGGATTTCCCTCTGTGTTGGAACTAAACCAATTTACTATATGT 2570
Db 2569 TGGCAACTTGATCAGTAAGGATTTCCCTCTGTGTTGTAACATAAACCACTACTATATGT 2628
QY 2571 TAGACAAGACATTTTTTTTTTTTTTCTCTCTCTGAAAAA-AAAATGAGGGAAGAGACA 2624
Db 2629 TAGACATGACATTTCTTTTCTCTCTCTGAAAAATAAAGTGTGGGAAGAGACA 2683

RESULT 10
AX677738

LOCUS AX677738 1683 bp DNA linear PAT 27-MAR-2003
DEFINITION Sequence 516 from Patent WO02086122.
ACCESSION AX677738
VERSION AX677738.1 GI:29335143
KEYWORDS

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1

REFERENCE 1
AUTHORS Legrain,P. and Daviet,L.
TITLE Protein-protein interactions in adipocytes
JOURNAL Patent: WO 02086122-A 516 31-OCT-2002;

FEATURES
Hybrigenics (FR)
Location/Qualifiers
source 1. .1683
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 62.8%; Score 1671; DB 6; Length 1683;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1682; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 60 ATGGAATGTCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 ATGGAATGTCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 60

QY 120 GCCGCCAAACGATTTTCATGATGTCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 61 GCCGCCAAACGATTTTCATGATGTCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 120
QY 180 CACAATCAATTAAATGGCTGGTCTTCTGTATGAAAAATGACTGGAATGAAAAACTCTACCCA 239

Db 121 CACAATCAATTAAATGGCTGTTCTCTGATGAAATGAATGGAATGAAATCTACCCA 180
QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAAAAAATCCTGGAAGGGAGGCCGTGTGAGGCG 299
Db 181 GTGTGGAAGCGGGAGACATGAGGTGGAAAAAATCCTGGAAGGGAGGCCGTGTGAGGCG 240
QY 300 GTCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTG 359
Db 241 GTCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTG 300
QY 360 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 301 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 360
QY 420 AGAAATGAGCGTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGGTCAGAG 479
Db 361 AGAAATGAGCGTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGGTCAGAG 420
QY 480 GACAGTGACGGGAAAAATGGACCGGCCAAAAGCCATATAACGTCTTCCCTGATGGGAAA 539
Db 421 GACAGTGACGGGAAAAATGGACCGGCCAAAAGCCATATAACGTCTTCCCTGATGGGAAA 480
QY 540 CCTTTTCTCACCACCCCGGATGGAGAAGATGGAATTTCACTACGTCTTCCACACACTT 599
Db 481 CCTTTTCTCACCACCCCGGATGGAGAAGATGGAATTTCACTACGTCTTCCACACACTT 540
QY 600 GGTCACTATTTCCAGAAATTTGGGACGATGTTTCACTGAGAGTTTCTGTGAACACAGCCAA 659
Db 541 GGTCACTATTTCCAGAAATTTGGGACGATGTTTCACTGAGAGTTTCTGTGAACACAGCCAA 600
QY 660 GTGACACTTGGGCCTCAACTCATGGAAGTGAATGCTGTCTACAGAAAGACATGGACGGGCATAT 719
Db 601 GTGACACTTGGGCCTCAACTCATGGAAGTGAATGCTGTCTACAGAAAGACATGGACGGGCATAT 660
QY 720 GTTCCCATCGCACAAAGTGAAAGATGTGTACGTGGTAACAGATCAGATTCCTGTGTTGTG 779
Db 661 GTTCCCATCGCACAAAGTGAAAGATGTGTACGTGGTAACAGATCAGATTCCTGTGTTGTG 720
QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTCCCTCAAGATCTCCCC 838
Db 721 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTCCCTCAAGATCTCCCC 780
QY 839 ATTATGTTTGATGTCCTGATTCATGATCCTPAGCCACTTCCTCAATTAATCTACCATTAAC 898
Db 781 ATTATGTTTGATGTCCTGATTCATGATCCTAGCCACTTCCTCAATTAATCTACCATTAAC 840
QY 899 TACAAGTGGAGTTTCGGGGATAATACTGGCCTGTTTGTGTTCCACCAATCATACTGTGAAT 958
Db 841 TACAAGTGGAGTTTCGGGGATAATACTGGCCTGTTTGTGTTCCACCAATCATACTGTGAAT 900
QY 959 CACACGTATGTGCTCAATGGAACCTTCAGCCCTTAACCTCAGTGTGAAGTGCAGCACCA 1018
Db 901 CACACGTATGTGCTCAATGGAACCTTCAGCCCTTAACCTCAGTGTGAAGTGCAGCACCA 960
QY 1019 GGACCTTGTCGGCCACCGCCACCAACCCAGACCTTCAAACCCCTTCTTTAGGA 1078
Db 961 GGACCTTGTCGGCCACCGCCACCAACCCAGACCTTCAAACCCCTTCTTTAGGA 1020
QY 1079 CCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTAAC 1138
Db 1021 CCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTAAC 1080
QY 1139 AGATATGGCCACTTTCAAGCCACCATCACAAATTPAGAGGGAATCTTAGAGGTTAAACATC 1198
Db 1081 AGATATGGCCACTTTCAAGCCACCATCACAAATTPAGAGGGAATCTTAGAGGTTAAACATC 1140
QY 1199 ATCCAGATGACAGACGTCCTGATGCCGTGCCATGGCTGAAAGCTCCCTAATAGACTTT 1258
Db 1141 ATCCAGATGACAGACGTCCTGATGCCGTGCCATGGCTGAAAGCTCCCTAATAGACTTT 1200
QY 1259 GTCGTGACCTGCCAAGGGAGCATTCGCCAGGAGGTCTGTACCATCATTTTCTGACCCACC 1318

Db 1201 GTCGTGACCTGCCAAGGGAGCAATTCACCGGAGGTCTGTACCATCATTTTCTGACCCACC 1260
QY 1319 TCGGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTGCTG 1378
Db 1261 TCGGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTGCTG 1320
QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCTGGGGAT 1438
Db 1321 ACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCTGGGGAT 1380
QY 1439 GACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCCTGACAGAGACCCAGCCTCG 1498
Db 1381 GACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCCTGACAGAGACCCAGCCTCG 1440
QY 1499 CCTTTAAGGATGGCAAAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACT 1558
Db 1441 CCTTTAAGGATGGCAAAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACT 1500
QY 1559 GTGATCTCCCTCTTGGTGTACAAAAAACACAAGGAATACACCCCAATAGAAAATAGTCCT 1618
Db 1501 GTGATCTCCCTCTTGGTGTACAAAAAACACAAGGAATACACCCCAATAGAAAATAGTCCT 1560
QY 1619 GGAATGTGGTCAGAAGCAAGGCCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTT 1678
Db 1561 GGAATGTGGTCAGAAGCAAGGCCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTT 1620
QY 1679 TTCCCGGAAACCCAGGAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db 1621 TTCCCGGAAACCCAGGAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAGGAGTTTCT 1680
QY 1739 TAA 1741
Db 1681 TAA 1683

RESULT 11

BC061725 2282 bp mRNA linear ROD 06-NOV-2003
LOCUS Rattus norvegicus glycoprotein (transmembrane) nmb, mRNA (cdna
DEFINITION clone MGC:72320 IMAGE:5598392), complete cds.
ACCESSION BC061725
VERSION BC061725.1 GI:38197643
KEYWORDS MGC.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 2282)
AUTHORS Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshikiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettaman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalls,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
MEDLINE 22388257
PUBMED 12477932
REFERENCE 2 (bases 1 to 2282)

AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Jeff Green/Paturu Kondaiah, NCI.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nhgri.nih.gov
Akhter,N., Ayèle,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 137 Row: c Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 18959233.
Location/Qualifiers
1. .2282
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="MGC:72320 IMAGE:5598392"
/tissue_type="Prostate, pool of NCI_CGAP_Pr30, 40, 41 and NCI_CGAP_Pr29, 39, 42"
/clone_lib="NCI_CGAP_Pr49"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
1. .2282
/gene="Gpnmb"
/db_xref="LocusID:113955"
/db_xref="RGD:71008"
58. .1776
/codon_start=1
/product="glycoprotein (transmembrane) nmb"
/protein_id="AAH61725.1"
/db_xref="GI:38197644"
/db_xref="LocusID:113955"
/translation="MESLGGVLVFLLLAAGLPLQAARFRDVLGHEQYPDHMRNNQLRGWSSDENEDQLYPWRRGEGRWKDSWEGRVQAALTSDSPALVGSNITFVNVLVFPRCKEDANGNIYERNCRSDLELASDPYVYNWTGADDEDWEDSTSQGHILRFPDGKPFPRPHGRKKNVFVFTLGLQYFQKLGRCSARVSINTVNLTVGPQVMEVIVFRRHGRAYIPISKVKDVVITDQIPIFVTMYQKDRNSSDETEFLRDLPIFFDVLIDPSSHFLNYSAISKWNFGDNTGLFVSNHNTLNHTYVLNGTFNFNLTVQTAVPGPCSPSPSSSTSPSPASSPSPSLTPSPSLMPTGHKSMELSDISNENCRINRYGYFRATITIVDGILEVNIQVADVPPIPTPDNLSLMDFIIVTKGATPEACTIISDPTCQIAQNRVCSFVAVDELCLLSVRRAFNGSTYCVNFTLGDASLALTSALISIPGKDLGSPLRVNGVLISIGCLAMFVTWVTILLYKKHKTYKPIGNCTRNVRVKGKLSVFLSHAKAPFSRGRDKDPLLQDKPWML"
844. .1014
/note="PKD; Region: Repeats in polycystic kidney disease 1 (PKD1) and other proteins"
/db_xref="CDD:smart00089"
ORIGIN
Query Match 39.5%; Score 1051.2; DB 10; Length 2282;
Best Local Similarity 74.1%; Pred. No. 3.8e-248;
Matches 1405; Conservative 0; Mismatches 443; Indels 48; Gaps 4;

QY	31	CTTGGTGCCTCGTCCGTGAGAAATT	CAGCATGGAATGTCTACTATTCTCTGGGATTTTC	90
Db	29	CTGGCTGAGGCGGGCGCTCCGAGT	CAGCATGGAAGTCTCTGCGGGTCTCTGGTATTTTC	88
QY	91	TGCTCCTGGTGCAGAAATTGCCACT	TGATGCCGCCAAACGATTTCAATGATGTCTGGGCA	150
Db	89	TGCTGCTGGTGCAGGACTGCCGCT	CACGGGCCAAGCGGTTCCGTGATGTCTGGGCC	148
QY	151	ATGAAAGACCTTCTGCTTACAT	GAGGAGCACAATCAATTAATGGCTGTCTTCTGATG	210
Db	149	ATGAGCAGTATCCGGATCACAT	GAGGAGAACCAATTACCTGGTGGTCTTTCAGATG	208
QY	211	AAAATGACTGGAATGAAAAACTCT	ACCCAGTGTGGAAGCGGGGAGACATGAGGTGAAAAA	270
Db	209	AAAATGAATGGATGAACAGCTGT	ATCCAGTGTGGAGGGGAGAGGCAGATGGAAGG	268
QY	271	ACTCCTGGAAGGAGGCCGTGTG	CAGCGGTCTGTACCCAGTCTCTGACCCCTCGTGG	330
Db	269	ACTCCTGGGAAGGAGGCCGTGTG	CAGCGGTCTGTACCCAGTCTCTGACCCCTCGTGG	328
QY	331	GCTCAATATAACATTTGCGGTGA	ACCTGATATTCCTAGATGCCAAAAAGGAGATGCCA	390
Db	329	GTCCAATATCACCTTCGTAGTGA	ACCTGGTGTTCCTCCAGATGCCAGAAAGATGCCA	388
QY	391	ATGCCAACATAGTCTATGAGA	AGAACTGCAGAAATGAGGCTGTTTATCTGCTGATCCAT	450
Db	389	ACGGCAATATCGTCTATGAGA	GAACCTGCAGAAATGAGGCTGTTTATCTGCTGATCCAT	448
QY	451	ATGTTTACAACTGGACAGCATTT	GCGGTGAACCTGATATTCCTAGATGCCAAAAAGGAGATGCCA	510
Db	449	ATGTCTACAACTGGACCAACAG	CGGCAGACGATGAGGACTGGGAAGACAGCACCCAGCCAAAG	508
QY	511	GCCATCATAACGTCTTCCCTG	ATGCGGAAACCTTTTCTCACCCACCCCGGATGGAGAAGAT	570
Db	509	GCCAGCACCTCAGGTTCCCGA	CGGGAAAGCCCTTCCCTCGCCCCCACGGACGGAAGAAAT	568
QY	571	GGAAATTCATCTACGTCTTCCA	CACACTTGGTCAGTAATTCAGAAATTCGGACGATGTT	630
Db	569	GGAACTTCGTCTACGTCTTCCA	CACACTTGGTCAGTAATTCAGAAATTCGGTCGCTGTT	628
QY	631	CAGTGAGAGTTTCTGTGAACA	CACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAGTGA	690
Db	629	CAGCACGAGTTTCTATAAA	CACAGTCAACTTGACAGTTGSCCCTCAGGTCATGGAAGTGA	688
QY	691	CTGTCTACAGAAGACATGGAC	GGGCATATGTTCCCATCGCACAAAGTGAAGATGTGTACG	750
Db	689	TTGTCTTTCGAAGACACGCG	CGGCATACATTCCCATCTCCAAAGTGAAGACGTTGATG	748
QY	751	TGTTAACAGATCAGATTCCTG	TGTTTGTGACTATGTTCCAGAAAGACGATCGAAATTCAT	810
Db	749	TGATAACAGATCAGATCCCTA	TATTCGTGACCATGTACAGAAAGATGACCCGGAATCGT	808
QY	811	CCGACGAAACCTTCC-CAA	AGATCTCCCCATTATGTTTGTGATGTCTCTGATTCATGATCCTA	869
Db	809	CTGATGAAACCTTCTCAG	AGACCTCCCCATTCTTCGATGTCTTCATTCACGATCCCA	868
QY	870	GCCACTTCTCAATTATTTAC	CATTAACTACAAGTGGAGCTTCGGGGATAAATACTGGCC	929
Db	869	GTCAATTCCTCAACTACTCT	GCCATTTCCTACAAGTGGAACTTTGGGGACACACTGGCC	928
QY	930	TGTTTGTTCACCAATCATAC	TGTGAATCACACGTTGCTCAATGGAACCTTCAGCC	989
Db	929	TGTTTGTCTCCAACAATCA	CACATTTGAATCACACGTTGCTCAATGGAACCTTCAACT	988
QY	990	TTAACTCACTGTGAAAGCTG	CAGCACCCAGGACCTTGTCCGCCACCGCCACCAC- ----	1043
Db	989	TTAACTCACCGTGCAAACTG	CAGTGCCTGGGACCATGCCCTCACCCACACTTCGCCTT	1048
QY	1044	-----	-----CACCCAGACCTTCAAAACCCACCC	1067
Db	1049	CTTCTTCGACTTCTCCTTC	CGCCTGCATCTTCGCCTTCAACACATTATCAACACCTAGTC	1108
QY	1068	CTTCTTTAGGACCTGCTGGT	GTGACAAACCCCTGGAGCTGAGTAGGATTCTCTGATGAAAACT	1127

Db 1109 CCTCTTTAATGCCTACTGGCCACAAATCCATGGAGCTGAGTGACATTTCCAATGAAACT 1168
QY 1128 GCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCACAATTTGTAGAGGAATCTTAG 1187
Db 1169 GCCGAATAAACAGATATGGTTACTTTCAGAGCCACCATCACAATTTGTAGATGGAATCCTAG 1228
QY 1188 AGGTTAACATCATCCAGATGACAGACGTCTCTGATGCCGGTGCCATGGCCCTGAAAGCTCCC 1247
Db 1229 AAGTCAACATCATCCAGGTAGCAGATGTCCCAATCCCAACACCGCAGCCTGACAACTCAC 1288
QY 1248 TAATAGACTTTGTGCTGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACCATCATTT 1307
Db 1289 TGATGACTTTCATTGTGACCTGCAAGGGGCCACTCCACGGAAGCCTGTACGATCATCT 1348
QY 1308 CTGACCCCACTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGA 1367
Db 1349 CTGACCCCACTGCGAGATCGCCCAAGAACAGGGTGTGAGCCCGGTGGCTGGATGAGC 1408
QY 1368 TGTGTCTGTGACTGTGAGACGAACCTTTCAATGGGTCTGGGACGTACTGTGTGAACCTCA 1427
Db 1409 TGTGCTCTCTGTCCGTGAGAGAGCCTTCAATGGGTCCGCACGTAATGTGTGAATTICA 1468
QY 1428 CCCTGGGGGATGACACAAAGCCTGGCTCTCACGAGCACCTGATTTCTGTTCCTGACAGAG 1487
Db 1469 CTCTGGGAGACGATGCAAGCCTGGCCCTCACGAGCCCTGATCTCTATCCCTGGCAAG 1528
QY 1488 ACCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCTTGATCTCCGTTGGCTGTGGCCA 1547
Db 1529 ACCTAGGCTCCCTCTGAGAACAGTAAATGGTGTCTGTATCCATTTGGCTGCCCGCCA 1588
QY 1548 TATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAAGGAATACAAACCCATAG 1607
Db 1589 TGTTGTCAACATGGTTACCATCTTGCTGTACAAAAACACAAGACGTACAAGCCATAG 1648
QY 1608 AAAATAGTCCTGGGAATGTGGTCAGAAAGCAAGGCCTGAGTGTCTTCTCAACCGTGCAA 1667
Db 1649 GAAACTGCACCAAGAACGTGGTCAAGGGCAAGGCCTGAGTGTCTTCTCAGCCATGCAA 1708
QY 1668 AAGCCGTGTTCTTCCCGGAAACCAAGGAAAGGATCCGCTACTCAAAAAACCAAGATTTA 1727
Db 1709 AAGCCCGTCTTCCCGAGGAGACCGGGAGAAAGGATCCACTGCTCCAGGACAAGCCATGGA 1768
QY 1728 AAGGAGTTCTTAAATTTTCGACCTTGTCTGAAGCTCACTTTTCAGTGCCATGTATGTG 1787
Db 1769 TGCTCTAAGTCTTCACTCTCACTTCTGACTGGGAACCACTCTCTGTG-CATGTATGIG 1827
QY 1788 AGATGCTGTGGAG---TGGCTATTAACTTTTTCCTAAAGATTATTGTTAAATAGAT 1843
Db 1828 AGCTGTGCAGAAGTACATGACTGGTAGTGTGTTTCTACGGATTATTGTAATATGTAT 1887
QY 1844 ATTGTGTTTGGGGAAGTTGAATTTTATAGTTA 1879
Db 1888 ATCATGGTTTAGGAGCGTAGTTAATTGGCATTTTA 1923

RESULT 12
BD062749
LOCUS BD062749 2303 bp DNA linear PAT 27-AUG-2002
DEFINITION Modulators of tissue regeneration.
ACCESSION BD062749
VERSION BD062749.1 GI:22608352
KEYWORDS JP 2001505761-A/3.
SOURCE Saccharomyces cerevisiae (baker's yeast)
ORGANISM Saccharomyces cerevisiae
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE 1 (bases 1 to 2303)
AUTHORS Nadel,M.S., Bonventre,J.V., Hession,C.A., Ichimura,T., Wei,H. and Cate,R.L.
Modulators of tissue regeneration
TITLE Patent: JP 2001505761-A 3 08-MAY-2001;
JOURNAL BIOGEN INC

COMMENT PN JP 2001505761-A/3
PD 08-MAY-2001
PF 23-MAY-1997 JP 1997542986
PR 24-MAY-1996 US 60/018228,23-AUG-1996 US 60/023442 PI
MICHELE SANICOLA NADEL,JOSEPH V BONVENTRE,CATHERINE A HESSION, PI
TAKAHARU ICHIMURA,HENRY WEI,RICHARD L CATE
PC C12N15/12,C07K14/47,C12N15/62,C07K16/18,A61K38/16,G01N33/50,
PC C12Q1/68,
PC C12N1/21,C12N5/10,C12N5/12,A61K48/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 107..1822.
FEATURES Location/Qualifiers
source 1..2303
/organism="Saccharomyces cerevisiae"
/mol_type="genomic DNA"
/db_xref="taxon:4932"
ORIGIN
Query Match 39.5%; Score 1051.2; DB 6; Length 2303;
Best Local Similarity 74.1%; Pred. No. 3.8e-248;
Matches 1405; Conservative 0; Mismatches 443; Indels 48; Gaps 4;
QY 31 CTTGGTGCCTCGTCCGTGAGAATTCCAGCATGGATGTAATGTTCTACTATTTCCTGGGATTTC 90
Db 78 CTGGCTGAGGGCGGCGTCCGAGTCAGCATGGAAGTCTCTCGGGGTCTCGGTATTC 137
QY 91 TGCTCCTGGCTGCAAGATTGCCACTTGTATCCCGCCCAACAGATTTCATGATGTCTGGCA 150
Db 138 TGCTGCTGGCTGCAGGACTCCGCTCCAGGCGGCCAAGCGGTTCGTGATGTCTGGGCC 197
QY 151 ATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGCTTCTGATG 210
Db 198 ATGAGCAGTATCCGGATCAGATGAGGGAGAACCAATTAACGTGGTGGTCTTCAGATG 257
QY 211 AAAATGACTGGAATGAAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGAAAA 270
Db 258 AAAATGAATGGATGAACAGCTGTATCCAGTGTGGAGGAGGGGAGGGCAGATGGAAGG 317
QY 271 ACTCCTGGAAGGGAGGCCGTGTGTCAGGGCGGTCTGTACCAGTGACTCACCAGCCCTCGTGG 330
Db 318 ACTCCTGGAAGGAGGCGCGTGTGCAGGCAGCCCTAACCCAGTATTACCCGGCCTTGGTGG 377
QY 331 GCTCAAAATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATGCCA 390
Db 378 GTTCAATATACCTTCGTAGTGAACCTGGTGTTCCTCCAGATGCCAGAAGGAAGATGCCA 437
QY 391 ATGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTATCTGTGATCCAT 450
Db 438 ACGCAATATCGTCTATGAGAGGAACCTGCAGAAATGATTTGGAGCTGGCTTCTGACCCGT 497
QY 451 ATGTTTACAACTGGACAGCATGTTCAGAGGACAGTACCGGGGAAAAATGGCACCAGCCAAA 510
Db 498 ATGTCTAACACTGGACCACAGGGGCAGACGATGAGGACTGGGAAGACAGCACCCAGCCAAG 557
QY 511 GCCATCAAAACGTCTTCCCTGATGGGAAACCTTTTCTCACCCACCCCGGATGGAGAAGAT 570
Db 558 GCCAGCACCTCAGGTTCCCGCAGCGGAAAGCCCTTCCCTCGCCCCACGGACGGAAGAAAT 617
QY 571 GGAATTTTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGATGTT 630
Db 618 GGAACCTTCGTCTACGTCTTCCACACACTTGGTCAGTATTTTCAAAAGCTGGGTGGTGT 677
QY 631 CAGTGAGAGTTTCTGTGAACACACAGCCAATGTGACACTTGGGCTCAACTCATGGAAGTGA 690
Db 678 CAGCACGAGTTTCTATAAAACACAGTCAAACTTGACAGTTGGCCCTCAGGTCATGGAAGTGA 737
QY 691 CTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGTACG 750
Db 738 TTGTCTTTTGAAGACACGGCGGGGCATACATTTCCCATCTCCAAAGTGAAGACGCTGTATG 797
QY 751 TGGTAACAGATCAGATTCCTGTGTTTGACTATGTTCCAGAAGAACGATCGAAATTCAT 810

Db 798 TGATAACAGATCAGATCCCTATATTCGTGACCATGTACCAGAAGAATGACCGGAACTCGT 857

QY 811 CCGACGAAACCTTCC-CAAAGATCTCCCATTTATGTTGATGTCTCTGATTCATCATCCTA 869

Db 858 CTGATGAAACCTTCCCTCAGAGACCTCCCAATTTCTTCGATGTCTCATTCACGATCCCA 917

QY 870 GCCACTTCCTCAATTATCTACCATTAACACTACAAGTGGAGCTTCGGGGATAAFACTGGCC 929

Db 918 GTCATTTCTCAACTACTCTGCCATTTCTTACAAGTGGAACTTTGGGACAACACTGGCC 977

QY 930 TGTGTTGTTCCACCAATCATACTGTGAATCACACGATGTCTCAATGGAACCTTCAGCC 989

Db 978 TGTGTTGTTCCCAACAATCACACTTTGAATCACACGATGTCTCAATGGAACCTTCAACT 1037

QY 990 TTAACCTCACTGTGAAAGCTGCAGCACCGAGGACCTTGTCCGCCACCGCCACCAC----- 1043

Db 1038 TTAACCTCACCGTGCRAAAGCTGCAGTCCGGGACCATGCCCTCACCCACACTTCGCCTT 1097

QY 1044 -----CACCCAGACCTTCAAAACCCACCC 1067

Db 1098 CTTCTTCGACTTCTCCTTCGCCTGCATCTTCGCCTTCACCCACATTTATCAACACCTAGTC 1157

QY 1068 CTTCTTTAGGACCTGCTGGTGACAAACCCCTCGAGCTGAGTAGGATTCCTGATGAAAACT 1127

Db 1158 CCTCTTTAATGCCTACTGGCCACAAATCCATGGAGCTGAGTGACATTTCCAATGAAAACT 1217

QY 1128 GCCAGATTAAAGATATGGCCACTTTTCAAGCCACCATCACAAATGTAGAGGGAATCTTAG 1187

Db 1218 GCCGAATAACAGATATGTTACTTCAGAGCCACCATCACAAATGTAGATGGAATCCTAG 1277

QY 1188 AGGTTAACATCATCCAGATGACAGACGTCCTGATCCCGTGCATGGCCTGAAAGCTCCC 1247

Db 1278 AAGTCAACATCATCCAGGTAGCAGATGTCCCAATCCCCACACCGCAGCTGCACAACTCAC 1337

QY 1248 TAATAGACTTTGTCGTGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACCATCATTT 1307

Db 1338 TGATGGACTTCATTTGACCTGCAAGAGGGCCACTCCCACGGAAGCCTGTACGATCATCT 1397

QY 1308 CTGACCCCACTGCGAGATCACCCAGAAACAGTCTGCAGCCCTGTGGATGTGGATGAGA 1367

Db 1398 CTGACCCCACTGCCAGATCGCCAGAACAGGGTGTGCAGCCCGTGGCTGTGGATGAGC 1457

QY 1368 TGTGTCGTGACTGTGAGACGAACTTCAATGGTCTGGACGPACTGTGTGAACCTCA 1427

Db 1458 TGTGCCCTCTGTCCGTGAGGAGAGCCTTCAATGGTCCGGCACGPACTGTGTGAATTTCA 1517

QY 1428 CCCTGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCTCTGACAGAG 1487

Db 1518 CTCCTGGAGACGATGCAAGCCTGGCCTCACCGCCCTGATCTCTATCCCTGGCAAAG 1577

QY 1488 ACCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCTTGATCTCCGTTGGCTGTGGCCA 1547

Db 1578 ACCTAGGCTCCCTCTGAGAACAGTGAATGGTGTCTCTGATCTCCATTTGGCTGTGGCCA 1637

QY 1548 TATTTGTCACCTGTGATCTCCCTCTTGGTGTAACAAAAACAAAGGAATACAAACCAATAG 1607

Db 1638 TGTTGTACCATGGTTACCATCTTGTGTACAAAAACAAAGACGTACAAGCCAATAG 1697

QY 1608 AAAATAGTCTGGGAATGTGTCAGAAAGCAAGGCCTGAGTGTCTTTCTCAACCGTGCAA 1667

Db 1698 GAAACTGCACCAAGAACGTGGTCAAGGGCAAGGCCTGAGTGTCTTCTCAGCCATGCAA 1757

QY 1668 AAGCCGTGTTCTTCCCGGAAACCAAGGAAAGGATCCGCTACTCAAAAACCAAGAAATTA 1727

Db 1758 AAGCCCGCTTCTCCCGAGGAGACCGGGAGAGGATCCACTGCTCCAGGACAAGCCATGGA 1817

QY 1728 AAGGAGTTTCTTAAATTTTCGACCTTGTCTTGAAGCTCACTTTTCAGTGCCATTGATGTG 1787

Db 1818 TGCTCTAAGTCTTCACTCTCACTCTCACTTCTGACTGGGAACCCACTTCTCTGTG-CATGATGTG 1876

QY 1788 AGATGTGCTGGAG-----TGGCTATTAAACCTTTTTTTCCTAAAGATTATTGTTAAATAGAT 1843

Db 1877 AGCTGTGCAGAAGTACATGACTGGTAGCTGTGTTGTTTCTACGGATTATTGTAATGTAT 1936

QY 1844 ATTGTGTTTGGGAAGTGAATTTTTTTTATAGGTTA 1879

Db 1937 ATCATGTTTAGGAGCGTAGTTAATTGGCATTTTA 1972

RESULT 13

AF322054

LOCUS

DEFINITION

AF322054

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

AF322054

Mus musculus dendritic cell-associated transmembrane protein precursor (Dchil) mRNA, complete cds.

AF322054.1

GI:13172896

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 2279)

Shikano,S., Bonkobara,M., Zukas,P.K. and Ariizumi,K.

Molecular Cloning of a Dendritic Cell-Associated Transmembrane Protein, DC-HIL, that Promotes RGD-Dependent Adhesion of Endothelial Cells Through Recognition of Heparan Sulfate Proteoglycans

Unpublished

2 (bases 1 to 2279)

Shikano,S. and Ariizumi,K.

Direct Submission

Submitted (16-NOV-2000) Dermatology, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75390-9069, USA

Location/Qualifiers

1. .2279

/organism="Mus musculus"

/mol_type="mRNA"

/strain="BALB/c"

/db_xref="taxon:10090"

/cell_line="XS52 DC"

/cell_type="dendritic cell"

1. .2279

/gene="Dchil"

44. .1768

/gene="Dchil"

/function="heparan sulfate proteoglycan-dependent integrin ligand"

/note="binds heparin; contains RGD motif; glycosylated; ITAM; similar to Mus musculus numb product encoded by GenBank Accession Number AJ251685"

/codon_start=1

/product="dendritic cell-associated transmembrane protein precursor"

/protein_id="AAK14240.1"

/db_xref="GI:13172897"

/translation="MESLCGVLGFLLLAAGLPQAARFRDVLGHEQYPNHRHNQLRGWSSDENEDHLYPVWRGDGRWKDSWEGRVQAVLTSDSPALVGSNITFVNVLVFPRCKEDANGNIYVEKNCRNDLGLTSDLHVYNWTAGDDGDWEDGTSRSQHLRFPPDRPFPRPHGWKKMSFVYVHTLGOYFQKLGRCARSINTVNLTAGPQVMEVTVFRRYGRAYIPISKVDVYITDQIPVFVTMSQKNDRLNLSDEIFLRDLPIVFVDVLHPSHFLNDSAISYKNWFGDNTGLFVSNHNTLNHTYVNLGTFNLTQVAVPGPCPPSPSTPPPPSTPPSPPLSTLSTPSPSLMPTGYKSMELSDISNENCRINRYGYFRATITIVEGILEVSIQIADVPMPTPOPANSLMDFTVTCKGATPMEACTIISDPTCQIAQNRVCSPVAVDGLCLLSVRRAFNGSGTYCVNFTLGDASLALTSTLISIPGKDPDSPRAVNGVLISIGCLAVLVTMTILLYKKHKAYKPIGNCPNTVKGLSVLLSHAKAPFRGDQEKDPLLQDKPRTL"

ORIGIN

Query Match 39.5%; Score 1049.8; DB 10; Length 2279;

Best Local Similarity 74.0%; Pred. No. 8.3e-248;

Matches 1409; Conservative 0; Mismatches 442; Indels 54; Gaps 4;

QY 30 CCTTGGTGCCTCCGTCCGTGAGAATTCAGCATGGAATGTCTCTACTATTTCCTGGGATTT 89

Db 14 CCTGACTGGTTGCAGCGCTCGGAGTCAGCATGGAAGTCTCTCGGGGCTCTGGGATTT 73
QY 90 CTGCTCCTGGCTGCAAGATTGCCACTTGATCCGCAAAACGATTTCATGATGCTGGGC 149
Db 74 CTGCTGCTGGCTGCAGACTGCTCTCCAGGCTGCCAAGCGATTTCGTGATGCTGGGC 133
QY 150 AATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGGTCTTCTGAT 209
Db 134 CATGAACAGTATCCCAATCACATGAGAGAGCACAAACCAATACGTGGCTGGTCTTCGGAT 193
QY 210 GAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGAAA 269
Db 194 GAAATGAATGGATGAACACCTGTATCCAGTGTGGAAGGGGAGACGGCAGGTGGAAG 253
QY 270 AACTCCTGGAAGGAGCGCTGTGCAGGCGGTCTTGACCAAGTCAACAGCCCTCGTG 329
Db 254 GACTCCTGGGAAGAGCGCGTGTGCAGGCAGTCTTGACCAAGTCAACCCGCTCTGGTG 313
QY 330 GGCTCAAAATATAACATTTGCGGTGAACCTGATATTCCTTAGATGCCAAAAGGAAGATGCC 389
Db 314 GGTTCCTCAATATCACCTTTGTGTGAACCTGGTGTTCCTCCAGATGCCAGAAGGATGCT 373
QY 390 AATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGCTGGTTTATCTGCTGATCCA 449
Db 374 AATGGCAATATCGTCTATGAGAAGAACTGCAGGAATGATTTGGACTGACCTCTGACCTG 433
QY 450 TATGTTTACAACCTGGACAGCATGGTCAGAGGACAGTACGGGAAAATGGCACCCGCCAA 509
Db 434 CATGTCTACAACCTGGACTGCAGGGCGAGATGATGGTGAATGGGAAGATGGCACCCAGCGA 493
QY 510 AGCCATCATAAACGTCTCCCTGATGGGAAACCTTTTCTCACACCCCGGATGGAGAAGA 569
Db 494 AGCCAGCATCTCAGGTTCCCGGACAGGAGGCCCTTCCCTCGCCCTCATGGATGGAAGAA 553
QY 570 TGGAAATTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGATGT 629
Db 554 TGGAGCTTTGTCTACGTCTTTCACACACTTGGCCAGTATTTCCAAAACCTGGGTCGGTGT 613
QY 630 TCAGTGAGAGTTTCTGTGAACACAGCAATGTGACACTTGGGCTCAACTCATGGAAGTG 689
Db 614 TCAGCACGGGTTTCTATAAACACACACTCAACTTGACAGCTGGCCCTCAGGTCAATGGAAGTG 673
QY 690 ACTGTCTACAGAAGACATGGAACGGGCATATGTTCCCATCGCACAGTGAAGATGTGTAC 749
Db 674 ACTGTCTTCGAAGATACGGCCGGGCATACATTTCCCATCTCGAAGGTGAAGATGTGTAT 733
QY 750 GTGGTAACAGATCAGATTCTCTGTGTTGTGACTATGTTCCAGAGAACGATCGAAATTC 809
Db 734 GTGATAACAGATCAGATCCCTGTATTCGTGACCATGTCCAGAGAAGATGACAGGAACCTTG 793
QY 810 TCCGACGAAACCTTCC-CAAAGATCTCCCAATATGTTTGTATGCTCTGATTCATGATCCT 868
Db 794 TCTGATGAGATCTTCTCAGAGACCTCCCATCGTCTTCGATGTCCTCATTCATGATCCC 853
QY 869 AGCCACTTCCTCAATATTCTACCAATTAACATAAGTGGAGCTTCGGGGATAAATCTGGC 928
Db 854 AGCCACTTCCTCAACGACTCTGCCAATTCCTACAAGTGAACCTTTGGGGACAACACTGGC 913
QY 929 CTGTTGTTTCCACCAATCATACTGTGAATCACACGATATGTGCTCAATGGAACCTTCAGC 988
Db 914 CTGTTTGTCTCAACAATCACACTTTGAATCACACTTATGTGCTCAATGGAACCTTCAAC 973
QY 989 CTTAACCTCACTGTGAAAGCTGCAGCACAGGACCTTGTCCGCCACCGC----- 1037
Db 974 CTTAACCTCACCGTGCAAACTGCAGTGCCTGGGCCATGCCCTCCCTTCGCTTCGACT 1033
QY 1038 -----CACCACCACCCAGACCTTCAAAA 1060
Db 1034 CCGCCTCCACCTTCAACTCCGCCCTCACCTCCGCCCTCACCTCTGCCACATATCAACA 1093
QY 1061 CCCACCCCTTTTAGGACCTGTGTTGACAACCCCTGGAGCTGAGTAGGATTCCTGAT 1120
Db 1094 CCTAGCCCCCTTTTAATGCCTACTGTGTTACAAATCCATGGAGCTGAGTGACATTTCCAAT 1153

QY 1121 GAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCACAAATTGTAGAGGA 1180
Db 1154 GAAAACTGCCGAATAAACAGATATGGCTACTTCAGAGCCACCATCACAAATTGTAGAGGG 1213
QY 1181 ATCTTAGAGGTTAAACATCATCCAGATGACAGACGCTCTGTGATGCCGTGCCATGGCCTGAA 1240
Db 1214 ATCCTGGAAGTCAGCATCATGCAGATAGCAGATGTCCCCATGCCACACCGCAGCCTGCC 1273
QY 1241 AGCTCCCTAATAGACTTTGTGCTGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACC 1300
Db 1274 AACTCCCTGATGGACTTCACTGTGACCTGCAAAAGGGGCCACCCCATGGAAGCCTGTACG 1333
QY 1301 ATCATTTCTGACCCCTCCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTG 1360
Db 1334 ATCATCTCCGACCCCTCCTGCGAGATCGCCAGAACCGGTCTGCAGCCCTGTGGCTGTG 1393
QY 1361 GATGAGATGTGCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGACGCTACTGTGTG 1420
Db 1394 GATGGCTGTGCCCTGCTGCTGTGAGAAGAGCCCTTCAATGGGTCTGGCACCTACTGTGTG 1453
QY 1421 AACCTCACCTTGGGGATGACACAAGCCTGGCTCTCAGAGCACCTCGATTTCTGTTCCT 1480
Db 1454 AATTTCACCTCTGGGAGATGATCAAGCCTGGCCCTCAGAGCACCTCGATCTCTATCCCT 1513
QY 1481 GACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGC 1540
Db 1514 GGCAAGACCCAGACTCCCTCTGAGAGCAGTGAATGGTGTCTCTGATCTCCATTTGGCTGC 1573
QY 1541 TTGGCCATATTTGTACATGTGATCTCCCTCTTGGTGTACAAAAACAAAGGAATACAAAC 1600
Db 1574 CTGGCTGTGCTTGTACCATGGTTACCATCTTGTCTGTACAAAAACAAAGGCGTACAAG 1633
QY 1601 CCAATAGAAAATAGTCTCTGGGAATGTGTGTCAGAAAGCAAGGCTGAGTCTTCTCAAC 1660
Db 1634 CCAATAGGAAACTGCCCCAGGAACACGGTCAAGGGCAAGGCTGAGTGTCTCTCCTCAGC 1693
QY 1661 CGTCAAAAAGCCGTGTTCTTCCCGGAAACCAAGAAAGGATCCGCTACTCAAAAACCAA 1720
Db 1694 CACGCAAAAGCCCGTTCTTCCGAGGAGACCAAGAGAGGATCCATTCCTCCAGGACAAG 1753
QY 1721 GAATTTAAAGGAGTTTCTTAAATTCGACCTTGTCTTGAAGCTCACTTTTCAGTGCCAT 1780
Db 1754 CCAAGCACACTCTAAGTCTTTGGCCTTCCCTCTGACCAGGAACCACTCTTCTGTG-CAT 1812
QY 1781 TGATGTGAGATGTGCTGGA----GTGGCTATTAAACCTTTTTCCTAAAGATTATTGTTA 1836
Db 1813 GTATGTGAGCTGTGCAGAAAGTATGTGGCTGGAACTGTTGTTCTCTAAGGATTATTGTAA 1872
QY 1837 AATAGATATGTGGTTTGGGGAAGTTGAATTTTATAGGTTAAA 1881
Db 1873 AATGTATATCGTGGCTTAGGGAGTGTGGTTAAATAGCATTTTATA 1917

RESULT 14
BC026375

LOCUS
DEFINITION

ACCESSION
VERSION

KEYWORDS
SOURCE

ORGANISM

REFERENCE

AUTHORS

BC026375 2299 bp mRNA linear ROD 06-OCT-2003
Mus musculus glycoprotein (transmembrane) nmb, mRNA (cDNA clone
MGC:31121 IMAGE:4164706), complete cds.
BC026375
BC026375.1 GI:20071338
MGC.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2299)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,

Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 2299)

Strausberg,R.

Direct Submission

Submitted (02-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305

Web site: <http://www-shgc.stanford.edu>

Contact: (Dickson, Mark) mcd@paxil.stanford.edu

Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

TITLE

JOURNAL
MEDLINE
PUBMED

REFERENCE
AUTHORS
TITLE
JOURNAL

REMARK
COMMENT

FEATURES
source

gene

CDS

misc_feature		GCLAVLVTMVTTILLYKHKAYKPIGNCPRNTVKGKLSVLLSHAKAPFFRGDQEKDPL LQDKPRIL"	
854. .1039		/note="PKD; Region: Repeats in polycystic kidney disease 1 (PKD1) and other proteins"	
/db_xref="CDD:smart00089"			
ORIGIN			
Query Match 39.5%; Score 1049.8; DB 10; Length 2299;			
Best Local Similarity 74.0%; Pred. No. 8.4e-248;			
Matches 1409; Conservative 0; Mismatches 442; Indels 54; Gaps 4;			
QY	30	CCTTGGTGCCTCGGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCTCTGGGATT	89
Db	50	CCTGACTGGTTGCAGGCGCTCGGAGTCAGCATGGAAAGTCTCTGCGGGTCTCTGGGATT	109
QY	90	CTGCTCCTGGCTGCAAGATTGCCACTTGATCCCGCCAAACGATTTTCATGATGTCTGGGC	149
Db	110	CTGCTGCTGGCTGCAGGACTGCCTCTCCAGGCTGCCAAGCGATTTCGTGATGTCTGGGC	169
QY	150	AATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGGTCTTCTGAT	209
Db	170	CATGAACAGTATCCCAATCACATGAGAGAGCACAAATACGTGGCTGGTCTTCCGAT	229
QY	210	GAATAATGACTGGAAATGAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGGA	269
Db	230	GAATAATGAATGGGATGAACACCTGTATCCAGTGTGAGGAGGGAGACGGCAGGTGGA	289
QY	270	AACTCCTGGAAGGAGGCGCGTGTGCAGGCGGTCTGACCACTGACTCACAGCCCTCGTG	329
Db	290	GACTCCTGGGAGGAGGCGCGTGTGCAGGCGTCTGACCACTGACTCACCGGCTCTGGTG	349
QY	330	GGCTCAAATATAACATTTGCGGTGAACCTGATATTTCCCTAGATGCCAAAGGAAGATGCC	389
Db	350	GGTCCAAATATCACCTTTGTGGTGAACCTGGTGTCCCGAGATGCCAGGAAGATGCT	409
QY	390	AATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTATCTGCTGATCCA	449
Db	410	AATGGCAATATCGTCTATGAGAAGAACTGCAGGAATGATTTGGGACTGACCTCTGACCTG	469
QY	450	TATGTTTACAACCTGGACAGCATGGTCAGAGGACAGTGACGGGGGAAATGGCACCGGCCAA	509
Db	470	CATGTCTACAACCTGGACTGCAGGGGCGAGATGATGCTGACTGGGAAGATGGCACCGCGA	529
QY	510	AGCCATCATAACGTCTTCCCTGATGGGAAACCTTTTCTCACACCCCGGATGGAGAAGA	569
Db	530	AGCCAGCATCTCAGGTTCCCGGACAGGAGGCCCTTCCCTCGCCCCATGGATGGAAGAA	589
QY	570	TGGAATTTTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGATGT	629
Db	590	TGGAGCTTGTCTACGTCTTTTCACACACTTGGCCAGTATTTCCAAAAAACTGGTCCGTGT	649
QY	630	TCAGTGAGAGTTTCTGTGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAGTG	689
Db	650	TCAGCAGGGTTTCTATAAACACAGTCAACTTGACAGCTGGCCCTCAGGTCAATGGAAGTG	709
QY	690	ACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAGTGAAGATGTGTAC	749
Db	710	ACTGTCTTTCGAAGATACGGCCGGGCATACATTTCCCATCTCGAAGGTGAAAGATGTGTAT	769
QY	750	GTGTAACAGATCAGATTCCCTGTGTTTGTGACTATGTTCCAGAAAGAACGATCGAAATCA	809
Db	770	GTGTAACAGATCAGATCCCTGTATTCTGTGACCATGTCCAGAAAGATGACAGGAACCTG	829
QY	810	TCCGACGAAACCTTCC-CAAAGATCTCCCCATTATGTTTGATGTCTCTGATTCATGATCCT	868
Db	830	TCTGATGAGATCTTCTCCTCAGAGACCTCCCCATCGTCTTCGATGTCTCTCATTCATGATCCC	889
QY	869	AGCCACTTCTCAATTATTCTACCATTAACACTACAAGTGGAGCTTCGGGGATAATACTGGC	928
Db	890	AGCCACTTCTCAACGACTCTGCCATTTCTCTACAAGTGGAACTTTGGGGACAACACTGGC	949
QY	929	CTGTTTGTTCACCAATCATACTGTGAATCACACGATATGTGCTCAATGGAACCTTCAGC	988

Db 950 CTGTTGTCTCCAACATCACACTTTGAATCACACTTATGTCTCAATGGAACCTTCAAC 1009
QY 989 CTTAACCTCACTGTGAAAGCTGCAGCACCAAGGACCTTGTCCGCCACCGC----- 1037
Db 1010 CTTAACCTCACTGTGAAAGCTGCAGTCCCGGCCATGCGCTCCCGCTTCCCGCTTCCGACT 1069
QY 1038 -----CACCACCAACCCAGACCTTCAAAA 1060
Db 1070 CCGCTCCACCTTCAACTCCGCCCTCACCTCCGCCCTCACCTCTGCCACATTAACAACA 1129
QY 1061 CCACACCTTCTTTAGGACCTGTGGTGACAAACCCCTGAGAGCTGAGTAGGATTCCTGAT 1120
Db 1130 CTTAGCCCTCTTTAATGCCTACTGGTTACAAATCCATGGAGCTGAGTGACATTTCCAAT 1189
QY 1121 GAAACTGCCAGATTAACAGATATGGCCACTTTCAAGCCACCATCAACAATTGTAGAGGA 1180
Db 1190 GAAACTGCCGAATAACAGATATGGCTACTTCAGAGCCACCATCAACAATTGTAGAGGG 1249
QY 1181 ATCTTAGAGTTAACATCATCCAGATGACAGACGTCTGATGCCGGTGCCATGGCCTGAA 1240
Db 1250 ATCTTGAAGTCAGCATCATGAGATAGAGATGTCCCCATGCCACACCCGACGCTGCC 1309
QY 1241 AGCTCCCTAATAGACTTTGTGCTGACCTGCCAAAGGGAGCATTCACACGGAGGTCTGTACC 1300
Db 1310 AACTCCCTGATGACTTCACTGTGACCTGCAAGGGGCCACCCCATGGAAGCCTGTACG 1369
QY 1301 ATCATTTCTGACCCACCTCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTG 1360
Db 1370 ATCATCTCCGACCCACCTGCCAGATCGCCAGAACCGGCTCTGCAGCCCTGTGGCTGTG 1429
QY 1361 GATGAGATGTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGAGACGCTACTGTGTG 1420
Db 1430 GATGGGCTGTGCTGCTGTGTGAGAGAGACCTTCAATGGGTCTGAGACCTACTGTGTG 1489
QY 1421 AACCTCACCTGGGGATGACACAAGCCTGGCTCTCAGGACACCCCTGATTTCTGTTCCCT 1480
Db 1490 AATTTCACTCTGGAGATGATGCAAGCCTGGCCCTCACAGACACCCCTGATCTATCCCT 1549
QY 1481 GACAGAGACCCAGCTCGCTTTAAGGATGGCAAAACAGTGCCTGATCTCCGTTGGCTGC 1540
Db 1550 GGCAAGACCCAGACTCCCTCTGAGAGCAGTGAATGGTGTCTCTGATCTCCATGGCTGC 1609
QY 1541 TTGGCCATATTTGTCACCTGTGATCTCCCTCTTGGTGACAAAAACACAGGAATACAAC 1600
Db 1610 CTGGCTGTGCTTGTACCATGTTTACCATCTTGTGTACAAAAACACAGGCGTACAAG 1669
QY 1601 CCAATAGAAAATAGTCTCTGGGAATGTGGTTCAGAAACAAAGGCTGAGTGTCTTCTCAAC 1660
Db 1670 CCAATAGGAACTGCCCCAGGAACACCGTCAAGGCAAAAGGCTGAGTGTCTCTCAGC 1729
QY 1661 CGTGCAAAAGCCGTGTTCTTCCCGGAAACAGGAAAAGGATCCGCTACTCAAAAACCAA 1720
Db 1730 CACGCGAAAGCCCGTCTTCCGAGGAGACCAGGAGAAGGATCCATTGCTCCAGACAAG 1789
QY 1721 GAATTTAAAGGAGTTTCTTAAATTTTCGACCTTGTTCGAAAGCTCACTTTTCAGTGCCAT 1780
Db 1790 CCAAGGACACTCTAAGTCTTTGGCCCTTCCCTCTGACAGGAACCCACTCTTCTGTG-CAT 1848
QY 1781 TGATGTGAGATGTGCTGGA---GTGGCTATTAACTTTTTCCTAAAGATTATGTTA 1836
Db 1849 GTATGTGAGCTGTGCAGAGATGTGGCTGGAACTGTGTTCTCTAAGGATTATGTAA 1908
QY 1837 AATAGATATTGTGGTTTGGGAAAGTTGAATTTTATAGGTTAAA 1881
Db 1909 AATGATATCGTGGCTTAGGGAGTGTGGTTAAATAGCATTTTAA 1953

RESULT 15
AR156839
LOCUS AR156839 2213 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 27 from patent US 6242419.
ACCESSION AR156839

VERSION AR156839.1 GI:15125543
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2213)
AUTHORS Strachan,L., Sleeman,M., Abernethy,N., Onrust,R., Kumble,A. and Murison,G.
TITLE Compositions isolated from stromal cells and methods for their use
JOURNAL Patent: US 6242419-A 27 05-JUN-2001;
FEATURES Location/Qualifiers
source 1..2213
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 39.4%; Score 1047.8; DB 6; Length 2213;
Best Local Similarity 74.1%; Pred. No. 2.6e-247;
Matches 1404; Conservative 0; Mismatches 437; Indels 54; Gaps 4;
QY 40 TCGTCCTCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTTCTCTGGGATTTCTGCTCCTGG 99
Db 3 TGCAGGCGCTCGAGTCAGCATGGAAGTCTCTCGGGGTCTCTGGGATTTCTGCTGCTGG 62
QY 100 CTGCAAGATTGCCACTTGTATGCCCCAAACGATTTTCATGATGTGCTGGCAATGAAAGAC 159
Db 63 CTGAGGACTGCCTCTCCAGGCTGCCAAGCGATTTCTGATGTGCTGGCCATGAACAGT 122
QY 160 CTTCTGCTTACATGAGGGAGCACAATCAATTAATTAATGGTGGTCTTCTGTGATGAAATGACT 219
Db 123 ATCCCAATCACATGAGAGAGCACAACCAATTAACGTGGTGGTCTTCGGATGAAATGAAT 182
QY 220 GGAATGAAAAAATCTACCCAGTGTGGAAGCGGGGAGACATGAGTGGAAAAAATCCTCTGGA 279
Db 183 GGGATGAACACTGTATCCAGTGTGGAGAGGGGAGACGGCAGGTGGAGGACTCCTGGG 242
QY 280 AGGAGGCGGTGTGCAGGCGGTCTGTACCAGTACTCACAGCCCTCTCTGGGCTCAAATA 339
Db 243 AAGGAGGCGGTGTGCAGGCGGTCTGTACCAGTACTCACCGGCTCTGTGGGTCCAATA 302
QY 340 TAACATTTGCGGTGAACCTGATATTTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACA 399
Db 303 TCACCTTTGTGGTGAACCTGGTGTTCCTCCAGATGCCAGAGGAAGATGTAATGGCAATA 362
QY 400 TAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACA 459
Db 363 TCGTCTATGAGAAGAACTGCAGGAATGATTTGGGACTGACCTCTGACCTGATGTTTACA 422
QY 460 ACTGGACAGCATGTGTGAGGAGCAGTGCAGGGGAAAAATGGCACCGGCCAAAGCCATCATA 519
Db 423 ACTGGACTGCAGGGGCAGATGATGGTGTGACTGGGAAGATGGCACCGCCGAAGCCAGCATC 482
QY 520 ACGTCTTCCCTGATGGGAAACCTTTTCTCTCACACCCCGGATGGAGAGATGGAATTTCA 579
Db 483 TCAGGTTCCCGGACAGGAGGCCCTTCCCTCGCCCCCATGGATGGAAGAAATGGAGCTTTG 542
QY 580 TCTACGTCTTCCACACACTTGGTCAAGTATTTCCAGAAAATTTGGACGATGTTCAAGTGAAG 639
Db 543 TCTACGTCTTTCACACACTTGGCCAGTATTTCCAAAACCTGGTTCGGTGTTCAGACCGGG 602
QY 640 TTTCTGTGAACACAGCCCAATGTGACACTTGGGCCCTCAACTCATGGAAGTGAATGTTTACA 699
Db 603 TTTCTATAAACACAGTCAACTTGCAGCTGGCCCTCAGGTCAATGGAAGTGAATGTTTTC 662
QY 700 GAAGACATGGACGGGCATATGTTTCCCATCGCACAGTGAAGATGTGTACGTGGTGAACAG 759
Db 663 GAAGATACGGCCGGGCATACATTTCCCATCTCGAAGGTGAAGATGTGTATGTGATAACAG 722
QY 760 ATCAGATTCCTGTGTTTGTGACTATGTTCCAGAACAGATCCGAAATTCATCCGACGAAA 819
Db 723 ATCAGATTCCTGTATTCGTGACCATGTCCAGAGAAATGACAGGAACTTGTCTGTATGAGA 782
QY 820 CCTTCC-CAAAGATCTCCCCCATATGTTTGTATGTCCTGATTCATGATCCTAGCCACTTCC 878

[illegible]

A;Title: Identification of a human melanoma antigen recognized by tumor-infiltrating lymphocytes
A;Reference number: A55753; MUID:94294401; PMID:8022805
A;Accession: A55753
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-161,'F',163-592,594-662 <KAW>
C;Keywords: glycoprotein

Alignment Scores:
Pred. No.: 6.05e-43 Length: 662
Score: 593.50 Matches: 167
Percent Similarity: 39.45% Conservative: 93
Best Local Similarity: 25.34% Mismatches: 200
Query Match: 12.24% Indels: 200
DB: 2 Gaps: 17

US-10-039-272-1 (1-2661) x I38400 (1-662)

QY	66	TGTCCTACTATTCTCGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCC	125
Db	8	CysLeu-----LeuHisLeuAlaValIleGlyAlaLeuLeuAlaValGlyAlaThr	24
QY	126	AAA-----CGATTTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGAG	179
Db	25	LysValProArgAsnGlnAspTrpLeuGlyValSerArg-----	37
QY	180	CACAATCAATTAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTACCCCA	239
Db	38	-----GlnLeuArg-----ThrLysAlaTrpAsnArgGlnLeuTyrPro	50
QY	240	GTGTGGAAGCGGGAGACATGAGGTGGAAAAAATCTCTGGAAGGAGGCCGTGTCAGGCG	299
Db	51	GluTrpThr-----GluAlaGlnArgLeuAspCysTrpArgGlyGlyGlnValSerLeu	68
QY	300	GTCCTGACCATGACTACACAGCCCTCGTGGGCTCAATATTAACATTTGCGGTGAACCTG	359
Db	69	LysValSerAsnAspGlyProThrLeuIleGlyAlaAsnAlaSerPheSerIleAlaLeu	88
QY	360	ATATTCCTTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC	419
Db	89	AsnPheProGlySerGlnLysValLeuProAspGlyGlnValIleTrpValAsnAsnThr	108
QY	420	AGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAAC-----TGG---	464
Db	109	-----IleIleAsnGlySerGlnValTrpGly	117
QY	465	-----ACAGCATGGTCAGAGACAGTGACGGGGAATAATGGCACCGCCAAAGCCATCAT	518
Db	118	GlyGlnProValTyrProGlnGluThrAspAla-----	129
QY	519	AACGTCTTCCCTGATGGGAAACCTTTTCCCTCACCCACCCCGGATGGAGAAGATGGAATTTC	578
Db	130	CysIlePheProAspGlyGlyProCysProSerGlySerTrpSerGlnLysArgSerPhe	149
QY	579	ATCTACGTCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGATGTTTCAGTCAGA	638
Db	150	ValTyrValTrpLysThrTrpGlyGlnTyrTrpGlnValLeuGlyGlyProValSerGly	169
QY	639	GTTTCTGTGAACACAGCCATGTGACACTTGGGCCTCAACTCATGGAAGTGAAGTGTCTAC	698
Db	170	LeuSerIleGlyThrGlyArgAlaMetLeuGlyThrHisThrMetGluValThrValTyr	189
QY	699	AGAAGACATGGA---CGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGTACGTGTA	755
Db	190	HisArgArgGlySerArgSerTyrValProLeuAlaHisSerSerAlaPheThrIle	209
QY	756	ACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAGAAGCATCGAAATTCATCCGAC	815
Db	210	ThrAspGlnValProPheSerValSerValSerGlnLeuArgAlaLeuAspGlyGlyAsn	229
QY	816	GAACCTTCCC-AAAGATCTCCCATTTATGTTTGATGTCCTGATTCATGATCCTAGCCAC	874
Db	230	LysHisPheLeuArgAsnGlnProLeuThrPheAlaLeuGlnLeuHisAspProSerGly	249

QY	875	TTCTCAATTAATCTACCATTAACATAAGTGGAGCTTCGGGATAATACTGSCCTGTTT	934
Db	250	TyrLeuAlaGluAlaAspLeuSerTyrThrTrpAspPheGlyAspSerSerGlyThrLeu	269
QY	935	GTTTCCACCAATCATCTGTGAATCACACAGTATGTCTCAATGGAACCTTCAGCCTTAAC	994
Db	270	IleSerArgAlaLeuValThrHisThrTyrLeuGluProGlyProValThrAlaGln	289
QY	995	CTCACTGTGAAAGCTGCAGACCA-----GGACCTTGTCCGCCACCGCCA---	1039
Db	290	ValValLeuGlnAlaAlaIleProLeuThrSerCysGlySerSerProValProGlyThr	309
QY	1039	-----	1039
Db	310	ThrAspGlyHisArgProThrAlaGluAlaProAsnThrThrAlaGlyGlnValProThr	329
QY	1039	-----	1039
Db	330	ThrGluValValGlyThrThrProGlyGlnAlaProThrAlaGluProSerGlyThrThr	349
QY	1039	-----	1039
Db	350	SerValGlnValProThrThrGluValIleSerThrAlaProValGlnMetProThrAla	369
QY	1039	-----	1039
Db	370	GluSerThrGlyMetThrProGluLysValProValSerGluValMetGlyThrThrLeu	389
QY	1039	-----	1039
Db	390	AlaGluMetSerThrProGluAlaThrGlyMetThrProAlaGluValSerIleValVal	409
QY	1039	-----	1039
Db	410	LeuSerGlyThrThrAlaAlaGlnValThrThrThrGluTrpValGluThrThrAlaArg	429
QY	1040	-----CCACCACCCAGACCTTCAAAACCC-----	1063
Db	430	GluLeuProIleProGluProGluGlyProAspAlaSerSerIleMetSerThrGluSer	449
QY	1064	---ACCCCTTCTTAGACCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGAT	1120
Db	450	IleThrGlySerLeuGlyProLeuLeuAspGlyThrAlaThrLeuArgLeuValLysArg	469
QY	1121	GAA-----AACTGCCAGATTAACAGATATGCCCACCTTTCAAGCCACCATCACAATT	1171
Db	470	GlnValProLeuAspCysValLeuTyrArgTyrGlySerPheSerValThrLeuAspIle	489
QY	1172	GTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATACAGACAGTCCTGATGCCGGTGCCA	1231
Db	490	ValGlnGly-----IleGluSerAlaGluIleLeuGlnAlaValPro	503
QY	1232	TGGCCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGCCAAGGGAGCATTCGCCAGGAG	1291
Db	504	SerGlyGluGlyAspAlaPheGluLeuThrValSerCysGlnGlyGlyLeuProLysGlu	523
QY	1292	GTCTGTACCATCATTTCTGACCCACCTGCGAGATCACCCAGAACACAGTCTGCAGCCT	1351
Db	524	AlaCysMetGluIleSerSerProGlyCysGlnProProAlaGlnArgLeuCysGlnPro	543
QY	1352	GTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGAAACCTTC---AATGGTCTGGG	1408
Db	544	ValLeuProSerProAlaCysGlnLeuValLeuHisGlnIleLeuLysGlyGlySerGly	563
QY	1409	ACGTACTGTGTGAACCTCACCTGGGGATGACACAGCCTGGCTCTCAGGAGCACCCCTG	1468
Db	564	ThrTyrCysLeuAsnValSerLeuAlaAspThrAsnSerLeuAlaValSerThrGln	583
QY	1469	ATTTCTGTTCTCGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAACAGTGCCTGATC	1528
Db	584	LeuIleMetProGlyGlnGluAlaGlyGly-----LeuGlyGlnValProLeuIle	600

QY 1529 TCCGTTGGCTGCTTGGCCATATTGTCTCACTGTGATCTCCCTCTTGGTGTACAAAAA 1585
Db 601 ValGlyIleLeuLeuValLeuMetAlaValLeuAlaSerLeuIleTyrArgArg 619
RESULT 4
A41234
melanocyte-specific protein Pmel-17 precursor - human
C;Species: Homo sapiens (man)
C;Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 30-Sep-1993
C;Accession: A41234
R;Kwon, B.S.; Chintamaneni, C.; Kozak, C.A.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.;
Proc. Natl. Acad. Sci. U.S.A. 88, 9228-9232, 1991
A;Title: A melanocyte-specific gene, Pmel 17, maps near the silver coat color locus on m
A;Reference number: A41234; MUID:92021023; PMID:1924386
A;Accession: A41234
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-668 <KWO>
A;Cross-references: GB:M77348
Alignment Scores:
Pred. No.: 1.65e-42 Length: 668
Score: 588.50 Matches: 170
Percent Similarity: 39.61% Conservative: 93
Best Local Similarity: 25.60% Mismatches: 198
Query Match: 12.13% Indels: 204
DB: 2 Gaps: 19
US-10-039-272-1 (1-2661) x A41234 (1-668)
QY 66 TGCTCTACTATTTCCTGGGATTTCTGCTCTGGCTGCAAGATTGCCACTTGATGCCGCC 125
Db 8 CysLeu-----LeuHisLeuAlaValIleGlyAlaLeuLeuAlaValGlyAlaThr 24
QY 126 AAA-----CGATTTCATGATGTGCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 25 LysValProArgAsnGlnAspTrpLeuGlyValSerArg----- 37
QY 180 CACAATCAATTAAATGGCTGGTCTTCTGATGAAAAATGACTGGAATGAAAAAACTCTACCCA 239
Db 38 -----GlnLeuArg-----ThrLysAlaTipAsnArgGlnLeuTyrPro 50
QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAATAAACTCTCTGGAAGGAGGCCGTGTGCAGCG 299
Db 51 GluTrpThr-----GluAlaGlnArgLeuAspCysTrpArgGlyGlyGlnValSerLeu 68
QY 300 GTCCTGACCAGTGACTCACAGCCCTCGTGGGCTCAAATATACATTTGCGGTGAACCTG 359
Db 69 LysValSerAsnAspGlyProThrLeuIleGlyAlaAsnAlaSerPheSerIleAlaLeu 88
QY 360 ATATTCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 89 AsnPheProGlySerGlnLysValLeuProAspGlyGlnValIleTrpValAsnAsnThr 108
QY 420 AGAAATGAGGCTGTTTATCTGCTGATCCCATATGTTTACAAC-----TGG--- 464
Db 109 -----IleIleAsnGlySerGlnValTrpGly 117
QY 465 -----ACAGCATGGTCAGAGGACAGTGACGGGAAATGGCACCGGCCAAAGCCATCAT 518
Db 118 GlyGlnProValTyrProGlnGluThrAspAla----- 129
QY 519 AACGTCTTCCCTGATGGGAACCTTTTCTCACCACCCCGGATGGAGAAGATGGAATTTC 578
Db 130 CysIlePheProAspGlyGlyProCysProSerGlySerTrpSerGlnLysArgSerPhe 149
QY 579 ATCTACGTCTTCCACACACACTTGGTCAGTATTTCCAGAAATTGGACGATGTTTCAGTGAGA 638
Db 150 ValTyrValTrpLysThrTrpGlyGlnTyrTrpGlnValLeuGlyGlyProValSerGly 169
QY 639 GTTCTGTGAACACAGCCAAATGTGACACTTGGGCCTCAACTCATGGAAAGTGACTGCTAC 698
Db 170 LeuSerIleGlyThrGlyArgAlaMetLeuGlyThrHisThrMetGluValThrValTyr 189

QY 699 AGAAGACATGGA---CGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTA 755
Db 190 HisArgArgGlySerArgSerTyrValProLeuAlaHisSerSerAlaPheThrIle 209
QY 756 ACAGATCAGATTCCCTGTGTTTGTGACTATGTTCCAGAAGAACGATCGAAATTCATCCGAC 815
Db 210 ThrAspGlnValProPheSerValSerValSerGlnLeuArgAlaLeuAspGlyGlyAsn 229
QY 816 GAAACCTTCCC-AAAGATCTCCCATTTATGTTTTCATGTCCTGATTCATGATCCTAGCCAC 874
Db 230 LysHisPheLeuArgAsnGlnProLeuThrPheAlaLeuGlnLeuHisAspProSerGly 249
QY 875 TTCCTCAATTATTCTACCATTAACACAGTGGAGCTTCGGGATAATACTGGCCTGTTT 934
Db 250 TyrLeuAlaGluAlaAspLeuSerTyrThrTrpAspPheGlyAspSerSerGlyThrLeu 269
QY 935 GTTTCACCAATCATACTGTGAATCACACGATGTCTCAATGGAACCTTCAGCCTTAAC 994
Db 270 IleSerArgAlaProValValThrHisThrTyrLeuGluProGlyProValThrAlaGln 289
QY 995 CTCACTGTGAAAGCTGCAGACCA-----GGACCTTGTCCGCCACCGCCA--- 1039
Db 290 ValValLeuGlnAlaAlaIleProLeuThrSerCysGlySerSerProValProGlyThr 309
QY 1039 ----- 1039
Db 310 ThrAspGlyHisArgProThrAlaGluAlaProAsnThrThrAlaGlyGlnValProThr 329
QY 1039 ----- 1039
Db 330 ThrGluValValGlyThrThrProGlyGlnAlaProThrAlaGluProSerGlyThrThr 349
QY 1039 ----- 1039
Db 350 SerValGlnValProThrThrGluValIleSerThrAlaProValGlnMetProThrAla 369
QY 1039 ----- 1039
Db 370 GluSerThrGlyMetThrProGluLysValProValSerGluValMetGlyThrThrLeu 389
QY 1039 ----- 1039
Db 390 AlaGluMetSerThrProGluAlaThrGlyMetThrProAlaGluValSerIleValVal 409
QY 1039 ----- 1039
Db 410 LeuSerGlyThrThrAlaAlaGlnValThrThrThrGluTrpValGluThrThrAlaArg 429
QY 1040 -----CCACCACCCAGACCTTCAAAAACCC----- 1063
Db 430 GluLeuProIleProGluProGluGlyProAspAlaSerSerIleMetSerThrGluSer 449
QY 1064 ---ACCCCTTCTTTAGGACCTGCTGGTGACAAACCCCTCGAGCTGAGTAGGATTCCTGAT 1120
Db 450 IleThrGlySerLeuGlyProLeuLeuAspGlyThrAlaThrLeuArgLeuValLysArg 469
QY 1121 GAA-----AAGTCCAGATTAAACAGATATGCCACTTTCAGCCACCATCACAATT 1171
Db 470 GlnValProLeuAspCysValLeuTyrArgTyrGlySerPheSerValThrLeuAspile 489
QY 1172 GTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGTGCCA 1231
Db 490 ValGlnGly-----IleGluSerAlaGluIleLeuGlnAlaValPro 503
QY 1232 TGGCCTGAAAAGCTCCCTAATAGACTTTTGTGCTGACCTGCCAAGGGAGCATTCACCGGAG 1291
Db 504 SerGlyGluGlyAspAlaPheGluLeuThrValSerCysGlnGlyGlyLeuProLysGlu 523
QY 1292 GTCTGTACCATCATTTCTGACCCCACTCGCGAGATCACCAGAACACAGTCTGCAGCCCT 1351
Db 524 AlaCysMetGluIleSerSerProGlyCysGlnProProAlaGlnArgLeuCysGlnPro 543

QY 1352 GTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGAAACCTTC---AATGGGTCTGGG 1408
Db 544 ValLeuProSerProAlaCysGlnLeuValLeuHisGlnIleLeuLysGlySerGly 563
QY 1409 ACGTACTGTGTGAACCTCACCTGGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTG 1468
Db 564 ThrTyrCysLeuAsnValSerLeuAlaAspThrAsnSerLeuAlaValValSerThrGln 583
QY 1469 ATTTCTGTTCCTGACAGAGACCCAGCCTCGCTTAAAGGATGGCAAACAGTGCCTG--- 1525
Db 584 LeuIleMetPro-----ValProGlyIleLeuLeuThrGlyGlnGluAlaGlyLeuGly 601
QY 1526 -----ATCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATC---TCCCTCTTG 1573
Db 602 GlnValProLeuIleValGlyIleLeuLeuValLeuMetAlaValValLeuAlaSerLeu 621
QY 1574 GTGTACAAAAA 1585
Db 622 IleTyrArgArg 625

RESULT 5

A49179
melanoma antigen homolog rpe1 - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 08-Nov-1996
C;Accession: A49179; I45861
R;Kim, R.Y.; Wistow, G.J.
Exp. Eye Res. 55, 657-662, 1992
A;Title: The cDNA RPE1 and monoclonal antibody HMB-50 define gene products preferentiall
A;Reference number: A49179; MUID:93122163; PMID:1478275
A;Accession: A49179
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-491 <KIM>
A;Experimental source: retinal pigment epithelium
A;Note: sequence extracted from NCBI backbone (NCBIN:122438, NCBIIP:122439)
C;Genetics:
A;Gene: RPE1

Alignment Scores:
Pred. No.: 2.38e-29 Length: 491
Score: 436.50 Matches: 124
Percent Similarity: 40.22% Conservative: 63
Best Local Similarity: 26.67% Mismatches: 123
Query Match: 9.00% Indels: 155
DB: 2 Gaps: 14

US-10-039-272-1 (1-2661) x A49179 (1-491)

QY 603 CAGTATTTCAGAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTG 662
Db 1 GlnTyrTrpGlnValLeuGlyGlyProValSerGlyLeuSerIleGlyThrAspLysAla 20
QY 663 ACATTGGGCCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGA---CGGGCATAT 719
Db 21 MetLeuGlyThrTyrAsnMetGluValThrValTyrHisArgArgGlySerGlnSerTyr 40
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTTGTG 779
Db 41 ValProLeuAlaHisSerSerSerAlaPheThrIleThrAspGlnValPropheSerVal 60
QY 780 ACTATGTTCCAG-----AAGAACGATCGAAATTCTATCCGACGAAACC 821
Db 61 SerValSerGlnLeuGlnAlaLeuAspGlyArgAsnLysArg--PheLeuArgLysGln- 79
QY 822 TTCCCAAAGATCTCCCATATATGTTTGTATGTCCTGATTCATGATCCCTAGCACTTCCTCA 881
Db 80 -----ProLeuThrPheAlaLeuGlnLeuHisAspProSerGlyTyrLeuA 95
QY 882 ATTATTCTACCATTAACATAAAGTGGAGCTTCGGGGATAATACTGGCCCTGTTGTTTCCA 941
Db 95 laGlyAlaAspLeuSerTyrThrTrpAspPheGlyAspSerThrGlyThrLeuIleSera 115

QY 942 CCAATCATCTGTGAATCACACGTATGTGCTCAATGGAACCTTCAGCCCTTAACCTCACTG 1001
Db 115 rgAlaLeuThrValThrHisThrTyrLeuGluSerGlyProValThrAlaGlnValValL 135
QY 1002 TGAAGCTGCAGCACCA-----GGACCTTGTCCGCCACCGCCACCACCAACCCA 1049
Db 135 euGlnAlaAlaIleProLeuThrSerCysGlySerSerProValProGlyThrThrAspA 155
QY 1050 GA----- 1051
Db 155 rgHisValThrThrAlaGluAlaProGlyThrThrAlaGlyGlnValProThrThrGluV 175
QY 1051 ----- 1051
Db 175 alMetGlyThrThrProGlyGlnValProThrAlaGluAlaProGlyThrThrValGlyT 195
QY 1051 ----- 1051
Db 195 rpValProThrThrGluAspValGlyThrThrProGluGlnValAlaThrSerLysValL 215
QY 1051 ----- 1051
Db 215 euSerThrThrProValGluMetProThrAlaLysAlaThrGlyArgThrProGluValS 235
QY 1052 -----CCTTCAAAACCCACCCCTTCTTTAGGA----- 1078
Db 235 erThrThrGluProSerGlyThrThrValThrGlnGlyThrThrProGluLeuValGluT 255
QY 1079 -----CCTGCTGTGCACAAC-----C 1094
Db 255 hrThrAlaGlyGluValSerThrProGluProAlaGlySerAsnThrSerSerPheMetP 275
QY 1095 CCCTGGAG-----CTGAGTAGGATTCTCTGATGAA----- 1123
Db 275 roThrGluGlyThrAlaGlySerLeuSerProLeuProAspThrAlaThrLeuValL 295
QY 1124 -----AACTGCCAGATTAAACAGATATGGCACTTTTCAAGCCA 1160
Db 295 euGluLysArgGlnAlaProLeuAspCysValLeuTyrArgTyrGlySerPheSerLeuT 315
QY 1161 CCATCACAATTGTAGAGGGAACTTTAGAGGTTAACATCATCCAGATGACAGACGTCCTGA 1220
Db 315 hrLeuAspIleValSer-----IleGluSerAlaGluIleLeuG 328
QY 1221 TGCCGCTGCCATGGCCTGAAAGTCCCTAAATAGACTTTGTCTGACCTGCCAAGGGAGCA 1280
Db 328 lnAlaValSerSerSerGluGlyAspAlaPheGluLeuThrValSerCysGlnGlyGlyL 348
QY 1281 TTCCACACGAGGTCTGTACCATCATTTCTGACCCCACTGCGAGATCACCCAGAACACAG 1340
Db 348 euProLysGluAlaCysMetAspIleSerSerProGlyCysGlnLeuProAlaGlnArgL 368
QY 1341 TCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTGACTGTGAGACGAAACCTTC---A 1397
Db 368 euCysGlnProValProProSerProAlaCysGlnLeuValLeuHisGlnValLeuLysG 388
QY 1398 ATGGGTCTGGACGTACTGTGTGAACCTCACCCCTGGGGGATGACACAAGCCTGGCTCTCA 1457
Db 388 lyGlySerGlyThrTyrCysLeuAsnValSerLeuAlaAspAlaasnSerLeuAlaMetV 408
QY 1458 CGAGCACCCCTGATTTCTGTTCTCTGACAGAGACCCAGCCTCGCCTTAAAGGATGGCAAACA 1517
Db 408 alSerThrGlnLeuValMetProGlyGlnGlu-----AlaGlyLeuArgGlnAlaPro- 425
QY 1518 GTGCCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACT---GTGATCTCCCTCTTGG 1574
Db 426 -----LeuPheValGlyIleLeuLeuValLeuThrAlaLeuLeuAlaSerLeuI 443
QY 1575 TGTACAAAAA 1585
Db 443 leTyrArgArg 446

QY 177 GAGCAATCAATTAAT-----GGCTGG 200
Db GluPheAsnTyrLeuAsnThrAlaArgThrLeuGluLeuTyrGlyValGluPheHisTyr 223
QY 201 TCTTCTGATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGACGGGGAGACATG 260
Db AlaArgAspGlnSerAsn--AsnGluIleMetIleGlyValMetSerGlyGlyIleLeu 242
QY 261 AGGTGGAATACTCTGGAAGGAGGCGGTGTGCAGGCGTCTGACCGAGTCACTACCA 320
Db IleTyrLysAsn-----ArgValArgMetAsnThrPhePro 254
QY 321 GCCCTCGTGGCTCAAAATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAG 380
Db TrpLeuLysIleValLysIleSerPheLysCysLysGlnPhePheIleGlnLeuArgLys 274
QY 381 GAAGATGCCAATGGC-----AACATAGTCTATGAGAAAGAACTGC 419
Db GluLeuHisGluSerArgGluThrLeuLeuGlyPheAsnMetValAsnTyrArgAlaCys 294
QY 420 AGAAATGAGGCTGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG 479
Db LysAsn-----LeuTrpLysAlaCysValGlu 303
QY 480 GACAGTGACGGGAAATGGCACCGGCCAAAGCCATCAATAACGTCTTCCCTCATGGAAA 539
Db -----HisHisThrPhePheArgLeuAspArg 312
QY 540 CCTTTTCTCACCCCGGATGGAGAGATGGAATTTTCATCTACGTCTTCCACACACTT 599
Db ProLeuProProGlnLys-----AsnPhePheAlaHisTyrPheThrLeu 327
QY 600 GGTCAGTATTTCAGAAATTTGGACGATGTTTCAGTG----- 635
Db GlySerLysPheArgTyrCysGlyArgThrGluValGlnSerValGlnTyrGlyLysGlu 347
QY 636 -----AGAGTTTCTGTGAACACAGCCAATGTGACACTTGGGCGCTCAACTC 680
Db LysAlaAsnLysAspArgValPheAlaArgSerProSerLysProLeuAlaArgLysLeu 367
QY 681 ATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAACTGAAA 740
Db MetAspTrpGluValValSerArgAsn-----SerIleSer 379
QY 741 GATGTGTACGTGTAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAGAAGACGAT 800
Db AspAspArgLeuGluThrGlnSerLeuProSerArgSerProProGlyThrProAsnHis 399
QY 801 CGAAATTCATCCGA-----CGAAACCTTCCCAAGATCTCCCCATTATG 844
Db ArgAsnSerThrPheThrGlnGluGlyThrArgLeuArgProSerSerValGlyHisLeu 419
QY 845 TTTGATGCTCTGATTCTAT---GATCCTAGCCAC---TTCCTCAATTATTCTACCATTAAC 898
Db ValAspHisMetValHisThrSerProSerGluValPheValAsnGlnArg----- 436
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATACTGTGAAT 958
Db -----SerProSerSerThrGlnAla 443
QY 959 CACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGACCA 1018
Db AsnSerIleValLeuGluSerSerProSer-----GlnGluThrPro 457
QY 1019 GGACCTTGTCCGCCACCGCCACCAACCAACCCAGA----- 1051
Db GlyAspGlyLysProProAlaLeuProProLysGlnSerLysLysAsnSerTrpAsnGln 477
QY 1051 ----- 1051
Db IleHisTyrSerHisSerGlnGlnAspLeuGluSerHisIleAsnGluThrPheAspIle 497

QY 1052 -----CCTTCAAAACCCACCCCTTCTTTAGGACCTGCTGGTGACAAACCCCTGGAG 1102
Db ProSerSerProGluLysProThrProAsnGlyGlyIleProHisAspAsnLeuValLeu 517
QY 1103 CTGAGTAGGATTCCTGATGAAAACTGCCAGATTAACAGATATGGC 1147
Db IleArgMetLysProAspGluAsn-----GlyArgPheGly 529
RESULT 8
T18518
apolipoprotein(a) - western European hedgehog (fragment)
C;Species: Erinaceus europaeus (western European hedgehog)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 07-Dec-1999
C;Accession: T18518
R;Lawn, R.M.; Boonmark, N.W.; Schwartz, K.; Lindahl, G.E.; Wade, D.P.; Byrne, C.D.; Fong
J. Biol. Chem. 270, 24004-24009, 1995
A;Title: The recurring evolution of Lp(a): Insights from cloning of hedgehog apolipoprotein
A;Reference number: I46259; MUID:96025778; PMID:7592597
A;Accession: T18518
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-2869 <LAW>
A;Cross-references: EMBL:U33170; NID:g1046358; PID:g1046359; PIDN:AAC48522.1
A;Experimental source: liver
C;Comment: The lipoprotein Lp(a), a major inherited risk factor for atherosclerosis, contains an apolipoprotein(a).
Alignment Scores:
Pred. No.: 0.162 Length: 2869
Score: 118.00 Matches: 141
Percent Similarity: 29.39% Conservative: 70
Best Local Similarity: 19.64% Mismatches: 248
Query Match: 2.43% Indels: 259
DB: 2 Gaps: 31
US-10-039-272-1 (1-2661) x T18518 (1-2869)
QY 6 CGAGGGCCAGAGGAATAAGTTAACCTTGGTGCGTCCGTGAGAAATTCAGCATGGAA 65
Db 410 ArgAsnProAspGlyGluIleAlaProTrpCys-----TyrThrThr-As 424
QY 66 TGTCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCAC---TTGATGCC 122
Db 424 n---SerAlaValArgTrpGluTyrCysSer---IleProAlaCysGluSerLeuSerPr 442
QY 123 GCCAAACGATTTTCATGATGTGCT---GGGCAATGAAGACCTTCTGCTTACATGAGGAG 179
Db 442 OProThrGluProMetVal-ValProGlyGln-----CysLeuGluGlyThrG 458
QY 180 CACAATCAATTAATGGC-----TG 200
Db 458 lyGluAsnTyrArg-GlySerValAlaValThrValSerGlyHisThrCysGlnArgTrp 477
QY 201 -----TCTTCTGATGAAAAATGACTGGAATGAAAACTCTACCCAGTGTGGAACGG 251
Db 478 ArgGluGlnSerProHisSerHisSerTrpThrProGluAsnTyrProThrLysAsnLeu 497
QY 252 GGAGACATGAGGTGGAAAAAATCTCTGGAAGGAGGCGCGTGTGCAG----- 296
Db 498 AspGlyAsnTyrCysArgAsnThr-----GlyGlyGluValAlaProTrpCysTyrThr 515
QY 297 -----GCGGTCTCTGACCACTGACTCACCAGTCACTCAGCCCTC 326
Db 516 ThrAsnSerAlaValArgTrpGluTyrCysSerIleProAlaCysGluSerProThr--- 534
QY 327 GTGGCTCAAATATAACATTTCGGGTGAACCTGATATTCCTAGATGCCAAAAGAGAT 386
Db 535 -----ProProThrGluGlnGlnVal 541
QY 387 GCCAATGGCAACATAGTC-----TATGAGAAGAAGCTGCAGAAATGAGGCTGTTTATCT 440
Db 542 ValProGlyGlnCysLeuGluGlyThrGlyGluAsnTyrArgGlySerValAlaValThr 561

QY	441	GCTGATCCATATGTT---	TACAACTGGACAGCATGGT	CAGAGGACAGTACGGGAAAAT	497
Db	562	AlaSerGlyHisThrCysGlnGlnTrpArgAlaGlnSerProHisSer-----HisSer	579		
QY	498	GGCACCGGCCAAAGCCATCATACGTCCTTCCCTGATGGGAAACCTTTTCTCACCACCCC	557		
Db	580	ArgThrProGluAsnTyrPro-ThrLysAsnLeuValGlyAsn--TyrCysArgAsnPr	598		
QY	558	GGATGGAGAAGATGGAATTTTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAA	617		
Db	598	oAspGlyGluIleAlaProTrpCysTyrThrThrAsnSerAlaVal-----Ar	614		
QY	618	TTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACACAGCCCATGTGACACTTGGGCCTCAA	677		
Db	614	gTrpGlu-TyrCysSerIleProAlaCysGluSerSerSerProThrGluProMetV	634		
QY	678	CTCATGGAAGTGACTGTCTACAGAAGACA-----TGGACGGGCATATGTTCCCATC	728		
Db	634	alValProGlyGlnCysLeuGluGlyThrGlyGluAsnTyrArgGlySerValAlaValT	654		
QY	729	GCACAAGTGAAAGA--TGTGTACGTGGTAACAGATCAGATTCCTGTGTTTGTGACTATG	785		
Db	654	hrValSerGluHisThrCysGlnArgTrpArgGluGlnSerProHisSerHisSerTrpT	674		
QY	786	TTCCAGAAGAACGATCGAAATTTCATCCGACGAAACCTTCCCAAAGATCTCCCCATTATGT	845		
Db	674	hrProGlu-----	676		
QY	846	TTGATGTCCTGATTTCATGATCCTAGCCACTTCCTCAATTATTCTACCATTAACACAGT	905		
Db	677	-----AsnTyrProThrLysAsnLeuAspG	685		
QY	906	GGAGCTTCGGGGATAATACTAGCCTGTTTGT-----TCCACCAATCATA	950		
Db	685	lyAsnTyrCysArgAsnThrGlyGlyGluValAlaProTrpCysTyrThrThrAsnSerA	705		
QY	951	CTGTGAATCACCGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTG	1010		
Db	705	laValArgTrpGluTyr-----CysS	712		
QY	1011	CAGCACCAGGACCTTGTCCGCCACCGCCACCCACCCAGACCTTCAAAACCCACCCCTT	1070		
Db	712	erileProAlaCysGluSerProThrProProThrGluGlnValValProGlyGlnC	732		
QY	1071	CTTTAGGACCTGCTGGTGACAAC-----	1093		
Db	732	ysLeuGluGlyThrGlyGluAsnTyrArgGlySerValAlaValThrAlaSerGlyHisT	752		
QY	1094	-----CCCCTGGAGCTGAGTAGGATTCCTGATGAAAACT	1127		
Db	752	hrCysGlnGlnTrpArgAlaGlnSerProHisSerHisSerArgThrProGluAsnTyrP	772		
QY	1128	GCCAGATTAAACAGATATGGCCACTTTCAGCCACCACCATCACAAATTGTAGAGGGAATCTTAG	1187		
Db	772	roThrLysAsnLeuValGlyAsnTyrCysArgAsn-----	783		
QY	1188	AGGTTAACATCATCCAGATGACAGACGTCCTGATGCCGGTGCCATGGCCCTGAAAGCTCCC	1247		
Db	784	-----ProAspGlyGluI	788		
QY	1248	TAATAGACTTTGTCGTGACCTGCCAAGGAGCATTCCCACGGAGGTCTGACCATCATTT	1307		
Db	788	leAlaProTrpCysTyrThrThrAsnSerAlaValArgTrpGluTyrCysSerIle----	806		
QY	1308	CTGACCCCACTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGA	1367		
Db	807	-----ProAlaCysGlu--SerLeuSerProProThrGluProMetvalValProGlyG	824		
QY	1368	TGTGTCTGCTGACTGTGAGACGAACCTTCAATGGGTCT-----	1405		
Db	824	lnCysLeuGluGlyThrGlyGluAsnTyrArgGlySerValAlaValThrValSerGlyH	844		
QY	1405	-----	1405		

```

Db      844  isThrCysGlnArgTrpArgGluGlnSerProHisSerHisSerTrpThrProGluAsnT  864
QY      1406  -----GGGACGTACTGTGTGAACCTCACCCCTGGGGAT-----  1438
Db      864  yrProThrLysAsnLeuAspGlyAsnTyrCysArgAsn---ThrGlyGlyGluValAlap  883
QY      1439  -----GACACAAAGCCTGGCTCTCACGAGACACCCCTGATTTCTGTCTCCTGAC---A  1484
Db      883  roTrpCysTyrThrThrAsnSerAlaValArgTrpGluTyrCysSerIleProAlaCysG  903
QY      1485  GAGACCCAGCCTCGCCTTAAGGATGGCAACAGTGCCTGATCTCCGTG-----  1535
Db      903  luSerProThrProThrGluGln-GlnValValProAspGlnCysLeuGluGlyThr  922
QY      1536  -----GCTGCTTGGCCATATTGT-----  1554
Db      923  GlyGluAsnTyrArgGlySerValAlaValThrAlaSer-GlyHisThrCysGlnGlnTr  942
QY      1554  -----  1554
Db      942  pArgAlaGlnSerProHisSerHisSerArgThrProGluAsnTyrProThrLysAsnLe  962
QY      1555  -----CACTGTGATCTCCCTCTTGGTGTACAAAAACACAGGAATACACCCAA  1604
Db      962  uValGlyAsnTyrCysArgAsnProAspGlyGluIleAlaProTrpCysTyrThrThrAs  982
QY      1605  TAGAAAAATAGTCTGGGAATGTGGTCAAGCAAGGCGCTGAGTGTCTTTCTCAACCGTG  1664
Db      982  nSerAlaValArgTrpGluTyrCysSerIleProAlaCysGluSerLeuSerProProTh  1002
QY      1665  CAAAAGCGGTCTTCTCCGGGAACCCAGGAAAAGGATCC  1704
Db      1002  rGluProMetValValProGlyGlnCysLeuGluGlyThr  1015

RESULT 9
S21766
diHydroLipoamide S-acetyltransferase (EC 2.3.1.12) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Nov-1993 #sequence_revision 01-Feb-1999 #text_change 05-May-2000
C;Accession: S21766
R;Matuda, S.; Nakano, K.; Ohta, S.; Shimura, M.; Yamanaka, T.; Nakagawa, S.; Tit
Biochim. Biophys. Acta 1131, 114-118, 1992
A;Title: Molecular cloning of dihydrolipoamide acetyltransferase of the rat pyru
s.
A;Reference number: S21766; MUID:92256482; PMID:1581353
A;Accession: S21766
A;Molecule type: mRNA
A;Residues: 1-555 <MAT>
A;Experimental source: heart
C;Genetics:
A;Genome: nuclear
C;Complex: component E2 of pyruvate dehydrogenase complex
C;Function:
A;Description: catalyzes conversion of acetyl-CoA and dihydrolipoamide to S-acet
A;Pathway: pyruvate metabolism
C;Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homolog
C;Keywords: acetyl-CoA; acyltransferase; cardiac muscle; coenzyme A; heart; lipo
F;7-81/Domain: lipoyl/biotin-binding homology <LPB1>
F;133-207/Domain: lipoyl/biotin-binding homology <LPB2>
F;254-307/Domain: component E3 binding #status predicted
F;46,172/Binding site: lipoamide (lys) (covalent) #status predicted
F;528,532/Active site: His, Asp #status predicted

Alignment Scores:
Pred. No.: 0.146 Length: 555
Score: 116.50 Matches: 78
Percent Similarity: 34.90% Conservative: 41
Best Local Similarity: 22.87% Mismatches: 129
Query Match: 2.40% Indels: 93
DB: 2 Gaps: 15

```

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 07-Dec-1999									
C;Accession: S50820									
R;Scott, J.; Leeck, C.; Forney, J.									
Nucleic Acids Res. 22, 5079-5084, 1994									
A;Title: Analysis of the micronuclear B type surface protein gene in Paramecium tetraurelia									
A;Reference number: S50820; MUID:95098630; PMID:7800503									
A;Accession: S50820									
A;Status: preliminary; nucleic acid sequence not shown; translation not shown									
A;Molecule type: DNA									
A;Residues: 1-2395 <SCO>									
A;Cross-references: EMBL:U07603; NID:g467226; PIDN:AAA81947.1; PID:g467227									
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1994									
C;Genetics:									
A;Genetic code: SGC5									
A;Introns: 472/3; 1310/3; 1821/3									
C;Superfamily: G surface protein									
Alignment Scores:									
Pred. No.:		0.344		Length:		2395			
Score:		114.00		Matches:		107			
Percent Similarity:		30.88%		Conservative:		48			
Best Local Similarity:		21.31%		Mismatch:		165			
Query Match:		2.35%		Indels:		182			
DB:		1		Gaps:		28			
US-10-039-272-1 (1-2661) x S50820 (1-2395)									
QY	282	GGAGCGCGTGTGCAGCGCGTCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAAATATA	341						
Db	1255	GlyAsnAlaCysGlnAlaPheSerThrCysGlu-----AlaLeuThrGlySerAsnLeu	1272						
QY	342	ACATTTCGGGTGAACCTGATATTCCTAGATGCCAA-----	377						
Db	1273	ThrTrpThrIle-----CysGlnAlaPheSerThrThrCysSerVal	1286						
QY	378	AAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGTTTA	437						
Db	1287	LysArgAspGlyThrGlyCysValThrIleGlnSerAlaCys-----ThrGlyTyr	1303						
QY	438	TCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAGGACAGTACGGGAAAAAT	497						
Db	1304	ThrThrIleAlaAsnCysTyrArgSerThrAla-GlyTyrCysThrAlaAsnSerGlyAs	1323						
QY	498	GGCACC GGCCAAAGCCATCATACGCTCTCCCTG-----ATGGGAAACCT	542						
Db	1323	pSerAlaCysGlnAlaIleSerAlaSerThrThrCysGluGlnIleLysLeuGlySerSe	1343						
QY	543	TTTCCTCACCACCCCGGATGGAGAAGATGGAATTCATCTACGCTTCCACACACTGGT	602						
Db	1343	rPheAlaPhe-----AspAspThrLysCysAsnThrPheLysThrGlyCysIl	1359						
QY	603	CAGTATTTCCAGAAATTGGG-----ACGATGTTCA	632						
Db	1359	eAlaLeuGlyThrSer-GlyCysGlnThrLysThrCysAlaAsnLysThrThrPropheA	1379						
QY	633	GTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAGTCACT	692						
Db	1379	la-----HisSerAspCysAsnThrThrTyrLeuSerThrCysThrSerAsnA	1394						
QY	693	GTCTACAGAAG-----ACATGGACGGGCATATGTTCCCATCGCACAAGTG	737						
Db	1394	laValSerSerProThrAlaCysThrThrMetAlaAlaThrCysAlaSerLeuThrThrS	1414						
QY	738	AAAGATGTGTACGTGGTAACATCAGAT-----TCCTGTGTT----	775						
Db	1414	erThrCysValTyrAlaValGluGlyGluCysValValSerGlyThrSerCysValArgL	1434						
QY	776	-----TGTGACTATGTTCCAGAAGAACGATCGAAATTCATCCGACGAAACCTTCCCAAAG	830						
Db	1434	ysThrCysAspThrAlaSerAlaAspThrSer-----	1444						
QY	831	ATCTCCCCCATTATGTTTGTATGCTCTGATTTCATGATCTCCTAGCCAC-----T	875						

957	ATCACACGATATGTCTCAATGGAACCTTACGCTTAACCTCAGTGTGAAAGCTGCAGCAC	1016
213	IleAlaAlaPheAlaAspTyrArgProThrGluValThrSerLeu--LysProGlnAlaP	232
1017	CAGGACCTTGTCCGCCACCG-----CCACCACCACCAGACCTTCAAAACCCA	1064
232	roProProValProProProValAlaAlaValProProIleProGlnProLeuAlaProT	252
1065	CCCCTTCTTTAGGACCTGCTGGTGAC-----AACCCCTGGAGCTGA	1106
252	hrProSerAlaAlaProAlaGlyProLysGlyArgValPheValSerProLeuAlaLysL	272
1107	GTAGGATTCCTGATGAATAAAGTCCAGATTAACAGATATGGCCACTTTTCAAGCCACCATCA	1166
272	ysLeuAlaAlaGluLysGlyIleAspLeuThr-----GlnValLysGlyT	287
1167	CAATTGTAGAGGA--ATCTTAGAGGTTAAACATCATCATCAGATGACAGCGTCTGTGATGC	1223
287	hrGlyProGluGlyArgIleIleLysLysAspIle-----AspSerPheValP	303
1224	CG-----GTGCCAT	1232
303	roThrLysAlaAlaProAlaAlaAlaAlaAlaProProGlyProArgValAlaProT	323
1233	GGCCTGAAAGCTCCCTAATAGACTTTTGTGCTGACCTTGCCAAAGGAGCATTTCCCACGGAGG	1292
323	hrProAlaGlyValPheIleAspIleProIle-----SerAsnI	336
1293	TCTGTACCATCATTTCTGACCCACCTGCGAGATCACCCAGAAACACAGTCTGCAGCCCT-	1351
336	leArgArgValIleAlaGln-----ArgLeuMetGlnSerLysGlnThrIleProH	353
1352	-----GTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGAACTTCA	1397
353	isTyrTyrLeuSerValAspValAsnMetGlyGluValLeuLeuValArgLysGluLeuA	373
1398	AT-----GGTCTGGGACGTACTGTGTGAACCTCACCTCGGGGATGACACAA	1445
373	snLysMetLeuGluGlyLysGlyLysIleSerValAsnAspPheIleIleLysAlaSerA	393
1446	GCCTGGCTCTCAGAGCACCTGATTTCTGTTCTCTGACAGACCCAGCC-----T	1496
393	laLeuAla-----CysLeuLysValProGluAlaAsnSerSerTrpMetAspT	409
1497	CGCCTTTAAGGATGGCAACAGTGCCTGTATCTCCGTT-----	1534
409	hrValIleArgGlnAsnHisValValAspValSerValAlaValSerThrProAlaGlyL	429
1535	-----GGCTGCTTGGCCATATTTGTCACGTG	1559
429	euIleThrProIleValPheAsnAlaHisIleLysGlyLeuGluThrIleAlaSerAspV	449
1560	TGATCTCCCTCTTGGTGTAACAAAAACAAGGAATACACCAATAGAAAATAGTCCCTG	1619
449	alValSerLeuAlaSerLysAlaArgGluGlyLysLeuGlnProHisGluPheGlnGlyG	469
1620	GGAATGTGGTCAGAAGC-----AAAGGCCTGAGTGTCTTTCTCA	1658
469	lyThrPheThrIleSerAsnLeuGlyMetPheGlyIleLysAsnPheSerAlaIleIleA	489
1659	ACCGTGCAAAAGCCGTGTTCTTCCCGGGAAACCAGGAAAGGATCCGCTACTCAAAAACC	1718
489	snProProGlnAlaCysIleLeuAlaIleGlyAlaSerGluAspLysLeuIleProAlaA	509
1719	AAGAATTTAAAGGAGTTTCTTAATTTCCAGCTTGTGTTTCTGAAGCTCACTTTTCAGTGC	1777
509	spAsnGluLysGlyPheAspValAlaSerVal-----MetSerValThrLeuSerCys	526

RESULT 10

550820

surface protein type 51B - Paramecium tetraurelia

;Species: Paramecium tetraurelia

Db 1445 -----PheAsp-----SerHisAlaGluCysIleAlaT 1454
QY 876 TCCTCAATTATTCTACCAT-----A 896
Db 1454 yrLeuSerThrCysThrValAlaArgThrGlyGlyCysGlnAlaArgAlaThrCysAlaS 1474
QY 897 ACTACAAG-----TGGAGCT 911
Db 1474 erTyrLysSerSerGlnGlnCysLysPheAsnSerThrGlyGlyLysCysPheTrpAsnP 1494
QY 912 TCGGGGAT-----AATACTGGCTGTTTGTTCACCAATCATCTG 953
Db 1494 roAsnAsnLysThrCysValAspLeuAsnCysGlyAsnIleGluAlaThrThrTyrA 1514
QY 954 TGAATCACACGTATGTCTCAATGGAACCTTCAGCCCTTAACCTCACTGTGAAA----- 1006
Db 1514 spThrHisAlaGluCysValAlaValAspThrGlnLeuLeuCysThrValArgAlaThrA 1534
QY 1007 -----GCTGCAGCACCAGGACCTTGTCCGCCACCGCCACCACCACCGACCTTCAAAAC 1061
Db 1534 snGlyAlaAlaValProGly----- 1540
QY 1062 CCACCCCTTCTTAGGACCTGTCTGGTGACAACCCCTTGAGCTGAGTAGGATTCCTGATG 1121
Db 1541 --CysMetAlaArgGlyAlaCysSerSerTyrSerIleGlu-----A 1554
QY 1122 AAAACTGCCAGATTAAACAGATATGGCCACTTTCAGGCCACCATCAACAATTGTAGAGGAA 1181
Db 1554 spGlnCysLysThrAsnProSerGly-----GlyV 1564
QY 1182 TCTTAGAGGTTAATCATCATCCAGATGACAGACGTCCTGATGCCGGTGCCATGGCCTGAAA 1241
Db 1564 alCysValTrpAsn-----ThrAsnLeuThrThrProValCysGlnAspLysS 1580
QY 1242 GCTCCCTAATAGACTTTGTCTGTGACC-----TGCCAAAGGG-----AGCA 1280
Db 1580 erCysThrThrAlaProThrAlaThrAlaThrHisAlaAspCysAspSerTyrPheSert 1600
QY 1281 TTCCACCGGAGGCTGTACCATCATTTCTTGACCCACC----- 1318
Db 1600 hrAlaThrIleLysCysThrValValAlaThrProAspThrAsnGlyGlyAlaAlaValL 1620
QY 1319 -----TGCGAGATCACCCAGAACACACAGTCTGCAGCCCTTGGATGTGGATGAGATGT 1370
Db 1620 euGlyGlyCysGln-----GlnThrAlaAlaCysSerSerTyrIleHisGlnGluGlnC 1638
QY 1371 GTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGT---GTGAACCTCA 1427
Db 1638 ys-----ArgPheAsnAlaThrGlyAspLeuCysGlyTrpAsnGlyT 1652
QY 1428 CCCTGGGGGATGACACAAGCCTGGCTCTCACGAGCACCTGATTTCTGTTCTCGACAGAG 1487
Db 1652 hrGlnCysAlaAspLysSerCysAlaThrAlaProAlaThrThrAspTyrAspAspAsnA 1672
QY 1488 AC 1489
Db 1672 sp 1672
RESULT 11
T02831
AAA protein L4171.3 [imported] - Leishmania major (strain Friedlin)
C;Species: Leishmania major
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 19-May-2000
C;Accession: C81460; T02831
R;Myler, P.J.; Audleman, L.; deVos, T.; Hixson, G.; Kiser, P.; Lemley, C.; Magness, C.;
Proc. Natl. Acad. Sci. U.S.A. 96, 2902-2906, 1999
A;Title: Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-c
A;Reference number: A81455; MUID:99178987; PMID:10077609
A;Accession: C81460
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1541 <PYL>
A;Cross-references: GB:AE001274; NID:g3264850; PIDN:AAC24654.1; PID:g2266909; GSPDB:GN00

A;Experimental source: strain MHOM/IL/81/Friedlin
C;Genetics:
A;Gene: L4171.3
A;Map position: 1

Alignment Scores:
Pred. No.: 0.377 Length: 1541
Score: 113.00 Matches: 46
Percent Similarity: 37.58% Conservative: 16
Best Local Similarity: 27.88% Mismatches: 60
Query Match: 2.33% Indels: 43
DB: 2 Gaps: 6

US-10-039-272-1 (1-2661) x T02831 (1-1541)

QY 1022 CCTTGTCGCCACCGCCACCACCCAGACCTTCAAAACCCACCCCTTCTTAGGACCT 1081
Db 156 ProProProProProProProProProProProProProProProProMetMetProProLeuAsnGly-Me 175
QY 1082 GCTGGTGACAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTAAACAGA 1141
Db 175 tilePhe-GlnProProGlyThr----- 182
QY 1142 TATGGCCACTTCAAGCCACCATCAACAATTGTAGAGGGAATCTTAGAGGTTAACATCATC 1201
Db 183 -----ProGlyMetProHisHisHis-----HisHisP 192
QY 1202 CAGATGACAGACGTCCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTTGTG 1261
Db 192 ro-----AlaGlyMetAlaThrProGln-----GlnHisLeuAlaS 204
QY 1262 GTGACCTGCCAAGGGAGGAGCAATCCACACGGAGGTCTGTACCATCATTTCTGACCCACCT-- 1319
Db 204 erGlnHisGlnHisGlnHisGlnHisGly--MetMetProMetTrpMetProProPr 223
QY 1320 -----GCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGT 1372
Db 223 oThrGlnThrGlnGlnHisGlnGlnGlnProGlyMetMetMetProMetAsnSerMe 243
QY 1373 CTGCTGACTGTGAGACGCAACCTTCAATGGGTCTGGAGCCTACTGTGTGAACCTCACCCCTG 1432
Db 243 tMetProMetSerMetProMetSerMetProMetGly-AlaAlaSerSerProLeuProp 263
QY 1433 GGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCTCTGACAGAGACCCA 1492
Db 263 roGlyGlnHisSerSerGlyProGlnGlnHisProPro-----GlnGlnProG 279
QY 1493 GCCTC 1497
Db 279 lyMet 280

RESULT 12
I55976
dihydrolipoamide S-acetyltransferase (EC 2.3.1.12), liver - rat (fragment)
N;Alternate names: primary biliary cirrhosis autoantigen
C;Species: Rattus norvegicus (Norway rat)
C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 05-May-2000
C;Accession: I55976; I71932
R;Gershwin, M.E.; Mackay, I.R.; Sturgess, A.; Coppel, R.L.
J. Immunol. 138, 3525-3531, 1987
A;Title: Identification and specificity of a cDNA encoding the 70 kd mitochondrial anti-
A;Reference number: I55976; MUID:87196380; PMID:3571977
A;Accession: I55976
A;Status: preliminary; translated from GB/EMBL/DBBJ
A;Molecule type: DNA
A;Residues: 1-457 <RES>
A;Cross-references: GB:D00092; NID:g220817; PIDN:AAA41813.1; PID:g220818; GB:M16075; NIJ
A;Experimental source: liver
C;Genetics:
A;Genome: nuclear
C;Complex: component E2 of pyruvate dehydrogenase complex
C;Function:
A;Description: catalyzes conversion of acetyl-CoA and dihydrolipoamide to S-acetyldihyd

A;Pathway: pyruvate metabolism
C;Superfamily: dihydrolipoamide acetyltransferase; lipoyl/biotin-binding homology
C;Keywords: acetyl-CoA; acyltransferase; coenzyme A; lipamide; mitochondrion
F;48-122/Domain: lipoyl/biotin-binding homology <LPB2>
F;169-222/Domain: component E3 binding #status predicted <E3B>

Alignment Scores:
Pred. No.: 0.309 Length: 457
Score: 112.50 Matches: 75
Percent Similarity: 34.97% Conservative: 39
Best Local Similarity: 23.01% Mismatches: 121
Query Match: 2.32% Indels: 91
DB: 2 Gaps: 14

US-10-039-272-1 (1-2661) x I55976 (1-457)

QY 957 ATCACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCAC 1016
Db 128 IleAlaAlaPheAlaAspTyrArgProThrGluValThrSerLeu--LysProGlnAlap 147
QY 1017 CAGGACCTTGTCGCCACCG-----CCACCACCACCCAGACCTTCAAAACCCA 1064
Db 147 roProProValProProProValAlaAlaValProProIleProGlnProLeuAlaProt 167
QY 1065 CCCCTTCTTTAGGACCTGCTGGTGAC-----AACCCCTCGAGCTGA 1106
Db 167 hrProSerAlaAlaProAlaGlyProLysGlyArgValPheValSerProLeuAlaLysL 187
QY 1107 GTAGGATTCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCA 1166
Db 187 ysLeuAlaAlaGluLysGlyIleAspLeuThr-----GlnValLysGlyT 202
QY 1167 CAATTGTAGAGGA---ATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGTATGC 1223
Db 202 hrGlyProGluGlyArgIleIleLysLysAspIle-----AspSerPheValP 218
QY 1224 CG-----GTGCCAT 1232
Db 218 roThrLysAlaAlaProAlaAlaAlaAlaProProGlyProArgValAlaProt 238
QY 1233 GGCCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGCCAAGGAGCATTTCCCACGGAGG 1292
Db 238 hrProAlaGlyValPheIleAspIleProIle-----SerAsnI 251
QY 1293 TCTGTACCATCATTTCTGACCCACCTGCGAGATCACCCAGAACACAGTCTGCAGCCCT- 1351
Db 251 leArgArgValIleAlaGln-----ArgLeuMetGlnSerLysGlnThrIleProH 268
QY 1352 -----GTGGATGTGGATGAGATGTGTCTGTCTGACTGTGTGAGACGAAACCTTCA 1397
Db 268 isTyrTyrLeuSerValAspValAsnMetGlyGluValLeuLeuValArgLysGluLeuA 288
QY 1398 AT-----GGGTCTGGGACGCTACTGTGTGAACCTCACCTCGGGGGATGACACAA 1445
Db 288 snLysMetLeuGluGlyLysGlyLysIleSerValAsnAspPheIleIleLysAlaSerA 308
QY 1446 GCCTGGCTCTCAGGACCCCTGATTTCTGTTCTGACAGAGACCCAGCC-----T 1496
Db 308 laLeuAla-----CysLeuLysValProGluAlaAsnSerSerTrpMetAspT 324
QY 1497 CGCCTTTAAGGATGGCAACAGTGGCCCTGATCTCCGTT----- 1534
Db 324 hrValIleArgGlnAsnHisValValaspValSerValAlavalserThrProAlaGlyL 344
QY 1535 -----GGCTGCTTGGCCATATTGTCACTG 1559
Db 344 euIleThrProIleValPheAsnAlaHisIleLysGlyLeuGluThrIleAlaSerAspV 364
QY 1560 TGATCTCCCTCTTGTGTACAAAAACACAGGAATACACCCCAATAGAAAATAGTCCTG 1619
Db 364 alValSerLeuAlaSerLysAlaArgGluGlyLysLeuGlnProHisGluPheGlnGlyG 384
QY 1620 GGAATGTGGTCAGAACG-----AAAGGCCTGAGTGTCTTTCTCA 1658

Db 384 lyThrPheThrIleSerAsnLeuGlyMetPheGlyIleLysAsnPheSerAlaIleIleA 404
QY 1659 ACCGTGCAAAAAGCCGTGTTCTTCCGGGAAAACAGGAAAGGATCCGCTACTCAAAAACC 1718
Db 404 snProProGlnAlaCysIleLeuAlaIleGlyAlaSerGluAspLysLeuIleProAlaA 424
QY 1719 AAGAAATTTAAAGGA 1732
Db 424 spAsnGluLysGly 428
RESULT 13
T15840
hypothetical protein C54G7.3 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C;Accession: T15840
R;Du, Z.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid C54G7.
A;Reference number: Z18416
A;Accession: T15840
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-2946 <DUZ>
A;Cross-references: EMBL:U40410; NID:g1065453; PID:g1065455; PIDN:AAA81392.1; CESP:C54G7.3
C;Genetics:
A;Gene: CESP:C54G7.3
A;Introns: 16/1; 53/2; 92/1; 160/3; 295/1; 346/1; 392/1; 440/1; 475/3; 579/1; 615/3; 1646/1; 2561/2; 2603/2; 2626/2; 2665/2; 2716/3; 2804/3; 2884/3
Alignment Scores:
Pred. No.: 0.597 Length: 2946
Score: 111.50 Matches: 126
Percent Similarity: 30.64% Conservative: 60
Best Local Similarity: 20.76% Mismatches: 201
Query Match: 2.30% Indels: 220
DB: 2 Gaps: 30
US-10-039-272-1 (1-2661) x T15840 (1-2946)
QY 472 GGTGAGGACAGTGACGGGGAAAATGGCACCGGCCAAAGCCATC-----ATA 519
Db 1101 GlyPheArgAlaThrSerGlyIleCysGluProAlaIleAlaValGlyGluProCysVal 1120
QY 520 ACGTCTTCC-----CTGATGGGAAAACCTTTTCTCACCACC 555
Db 1121 ThrSerAsnGlnCysPheAspGluSerGluCysValPheGlyIleCysThrCysThrGly 1140
QY 556 CCGATGGAGAAGAT-----GGATTTTCATCTACGTCTTCCACAC 597
Db 1141 ProAsnCysLysAspThrLysMetAlaHisProGlyGluAspCysThrSerLeuLysThr 1160
QY 598 TTGTCAGATATTTCCAGAAATTTGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCA 657
Db 1161 ValCysSer-TyrAsnSerTyrCysSerLeuMetSer---SerValCysGlu----- 1176
QY 658 ATGTGACACTTGGCCCTCAACTCATGGAAGTACTGTCTACAGAACACATGGACGGGCAT 717
Db 1177 -CysProSerGlyMetAlaThrLysGlyThrLysCys---GluAsnThrPheGluSerI 1195
QY 718 A-----TGTTCCTCATCGCACACTGAAAGATGTGTACGTGTGTAACAGATCAGATTC 768
Db 1195 eGlyLysAspCysVal-----ThrSerArgAsnCysGlnLysSerSerTyrCysAspAs 1213
QY 769 C-----TGTTTGTGAC-----TATGTTTCC 789
Db 1213 nGlyTyrCysValCysLysAsnGlyHisLysIleGlyGluAsnMetCysPheAsnSerPr 1233
QY 790 AGAAGAACGATCGAAATTC-----ATCCGACGAAACCTTCCCAAGATCTCCC 837
Db 1233 oSerGluTyrLysSerPheSerIleLeuProPheAspLysAsnIleGlyGlnAsnThrPr 1253

QY 838 C-----ATTATGTTTGATGTCCT 855
Db 1253 oLeuGlnAsnThrLeuLysAsnGluPheArgGlyLeuGlnGluIleSerAsnAspGlyLe 1273
QY 856 GATTCAATGAT----- 865
Db 1273 uPheHisThrSerThrLysTrpProGluIleLeuSerPheThrMetIleProProProPr 1293
QY 866 -----CCTAGCCACTTCCTC----- 880
Db 1293 oAspThrAsnLeuProAsnSerAsnLeuProGlnValPheSerSerPheProIleValTy 1313
QY 881 -----AATTATTCTACCAATTAACTACAAGTGGAGCTT 912
Db 1313 rGlyAlaLysThrValAlaGluGluAsnAsnSerThrMetLysTyrlLysIleAlaPh 1333
QY 913 CGGGGATAATACTGGCCTGTTTGTTCACCAATCATCTGTGAATCACACGTATGTGCT 972
Db 1333 e-----ProGlyGluTyrcysGlyThrGlyGlnValCysLeuGlyAsnSerValCy 1350
QY 973 CAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGACACGAGACCTTGTCCGCC 1032
Db 1350 sGluAsnGlnPheCysArgCysLeuGlnAspValAlaAlaGluAsnGlyIleCysProPr 1370
QY 1033 ACCGCCACCAACCAACCCAGACCTTCAAAACCCACCCCTTCCTTAGGACCTGCTGGTGACAA 1092
Db 1370 o-----GlnValAspAsnLeuArgValLeuGlyLeuGl 1381
QY 1093 CCCCCTG-----GAGCTGAG 1107
Db 1381 nProLeuGlyLysGluPheArgPheSerGluGlyLysLysIleGluMetArgArgThrSe 1401
QY 1108 TAGGATTCTGTATGAAAACTGCCAGATTAACAGATATGGCCACTTCAAGCCACCATCAC 1167
Db 1401 rSerLeuProLeuGluAsnCysGlnAsnGluGluValCysGluAsnSerThrCysGl 1421
QY 1168 AATTGTAGAGGGAATCTTAGAGTT----- 1192
Db 1421 nSerIleLeuGlyLeuGlyArgIleCysGlnCysValGluAsnThrValLeuTrpAsnGl 1441
QY 1193 -AACATCATCCAGATGACACACGTC-----CTGATCGCGTGCCA----- 1231
Db 1441 yAsnCysValIleValGluAspSerTyraSpLeuThrProIleAspGlyAsnCysAspGl 1461
QY 1232 -----TGGCCTGAAAGCTCCCTAATAGACTTTGTGCGTGACCTGCCAAGGGAG 1278
Db 1461 uAspSerMetCysLeuSerGlySerGluCysValAspGlyLysCysLeuCysSerAspGl 1481
QY 1279 C-----ATTCCCACGGAGGCTGTACCATCTTCTGACCCC-----ACCTGGCGAGAT 1326
Db 1481 yLysArgLeuIleLeuGlyIleCysValPheIleAlaLeuProGluThrSerCysGluAs 1501
QY 1327 CACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTCTGACTGTGAG 1386
Db 1501 nGlyGluValCysIleAsnGlySerValCysGlyAspSerAsnCysGluCysThr---Gl 1520
QY 1387 ACGAACCTTCAATGGGFTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGGATGACACAAG 1446
Db 1520 uAsnThrTyraSnHisAsnGlyAsn---CysValAspIleLysLeu-----AspGluSe 1537
QY 1447 CCTGGCTCTCACGACACCCCTGTATTCTGTCTTCGTACAGACACCCAGCCCTCGCCTTTAAG 1506
Db 1537 rLeuIleLeuArgGlnGluValGlyGlnAspSerAspGluLysAspSerGluSerLeu-- 1556
QY 1507 GATGGCAAACAGTGCCTGATCTCCGTTGGCTGCTGGCCATATTTGTCACTGTGATCTC 1566
Db 1556 ----- 1556
QY 1567 CCTCTTGGTGTAACAAAAACAAGGAATACACCCCAATAGAAAATAGTCTCTGGGAATGT 1626
Db 1557 -----SerGluAsnAsnAspAsnGluGluAsnSerArgSerLe 1569
QY 1627 GGTCAAGAACAGGCGCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGCTGTCTTCCCGGG 1686

Db 1569 uValArgArgGluLeuAlaSerIleAspCys-----AlaAs 1581
QY 1687 AAACCAAGGAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAGAGTTCCTTAAATTC 1746
Db 1581 nAspGlnGluCysGlnProAsnPheLysCysGlnGluTyraValCysValCysAspAsnSe 1601
QY 1747 GACC-----TTGTTTCTGAAGCTC----- 1765
Db 1601 rThrGluAsnCysLeuLysSerIleValAspLeuLysValSerValProProGlySerGl 1621
QY 1766 -----ACTTTTCAGTGCCATTGATGTGAGATGTGCTGGAGTGGCTATTAAACCTTT 1815
Db 1621 yCysSerGluThrArgLysCysGlyAspSerIleCysTyra----- 1635
QY 1816 TTTTCTAAAGATTATTGT 1834
Db 1636 -----LysAspTyrcys 1639
RESULT 14
KGHUGH
histidine-rich glycoprotein precursor - human
N;Alternate names: HRG
C;Species: Homo sapiens (man)
C;Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 16-Jun-2000
C;Accession: A01287; S29669
R;Koide, T.; Foster, D.; Yoshitake, S.; Davie, E.W.
Biochemistry 25, 2220-2225, 1986
A;Title: Amino acid sequence of human histidine-rich glycoprotein derived from the nucl.
A;Reference number: A01287; MUID:86216149; PMID:3011081
A;Accession: A01287
A;Molecule type: mRNA
A;Residues: 1-525 <KOI>
A;Cross-references: GB:AB005803; NID:g2280513; PIDN:BAA21613.1; PID:g2280514
R;Hennis, B.; Havelaar, A.; Klufft, C.
submitted to the EMBL Data Library, October 1991
A;Description: PCR detection of a dinucleotide repeat in the human histidine-rich glyco
A;Reference number: S29669
A;Accession: S29669
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 214-247 <HEN>
A;Cross-references: EMBL:Z17218; NID:g32453; PIDN:CAA78925.1; PID:g32454
C;Comment: Although its physiological function is not yet known, HRG does bind heme, dy
din, and the lysine-binding site of plasminogen. On the basis of its homology with HMW
lood coagulation cascade.
C;Comment: The amino half of this protein is homologous to the first two cystatin-like
ould not have inhibitory activity.
C;Comment: In addition to having a high histidine and proline content, this protein has
e-rich' region.
C;Genetics:
A;Gene: GDB:HRG
A;Cross-references: GDB:120055; OMIM:142640
A;Map position: 3q27-3q27
C;Superfamily: histidine-rich glycoprotein; cystatin homology
C;Keywords: duplication; glycoprotein; heparin binding; tandem repeat
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-525/Product: histidine-rich glycoprotein #status predicted <MAT>
F;19-131/Domain: cystatin homology <CY1>
F;140-246/Domain: cystatin homology <CY2>
F;276-321/Region: proline-rich
F;348-437/Region: histidine-rich
F;351-497/Region: proline-rich
F;63,125,344,345/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;78-89,105-126,218-241/Disulfide bonds: #status predicted
Alignment Scores:
Pred. No.: 0.431 Length: 525
Score: 111.00 Matches: 51
Percent Similarity: 29.50% Conservative: 8
Best Local Similarity: 25.50% Mismatches: 51
Query Match: 2.29% Indels: 91
DB: 1 Gaps: 13

US-10-039-272-1 (1-2661) x JQ0533 (1-1874)

QY	797	CGATCGAAATTCATCCGACGAAACCTTCCCAAAGATCTCCCATATTGTTTGATGTC---	853
Db	533	ArgMetSerPheThrThrValAsnLeuProLeuGlnProPrometTrpLeuAlaIleGly	552
QY	854	-----CTGATTCAATGATCCTAGCCACTTCTCTCAATTATTCTACCATTAACATAAGTGG	907
Db	553	AlaSerLeuValProGluLeuAlaPheLeuLeuSer-----	565
QY	908	AGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATACTGTGAATCACACGTAT	967
Db	566	LeuSerGlyAsp-----ValAspLeuGlnThrGlnHisAspIleTyrHisHis	582
QY	968	GTGCTCAATGGAACCTTCAGCCTTAACCTCACT-----GTGAAAGCTGCAGCACCA	1018
Db	583	LeuHisProGluAsnPheThrLeuSerTrpThrArgThrProTyrLeuAlaLeuAlaPro	602
QY	1019	GGACCT-----	1030
Db	603	SerProPheLeuProTyrAlaHisSerProLeuProProLeuProValAsnSerSerPro	622
QY	1031	-----CCACCGCCACACCCAGACCTTCAAAACCCACCCCTTCTTTAGGACCT---	1081
Db	623	LeuPheProProProProLeuProProSerGlnProProLeuSerGlnGlyProAla	642
QY	1081	-----	1081
Db	643	ThrGlnAlaProSerAlaGlnProThrProGlyGluProLeuLeuAlaProProThrThr	662
QY	1082	-----	1090
Db	663	GluLeuLysProGluSerSerAsnProAsnAsnProAsnProSerSerSerAlaGlySer	682
QY	1091	AACCCCTGGAGCTGAGTAGGATTCTCTGATGAAACTGCCAGATTAAACAGATATGGCCAC	1150
Db	683	AsnProProProLysSerSerSerSerAspAsnProAlaProAlaProAsnLysProThrPro	702
QY	1151	TTTCAAGCCACCATCACA-----	1168
Db	703	ThrSerSerSerThrThrProProSerProAsnLeuProLeuGlnPheGlySerIleHis	722
QY	1169	-----ATTGTAGAGGAATCTTAGAGGTTAACATCATC-----	1201
Db	723	SerProPheLeuSerAspGlyGlnLeuAsnTyrSerAlaLeuProProProGlnAspPro	742
QY	1202	---CAGATGACAGACGTCTGATGCCGTGCCATGGCCTGAAAGCTCCCTAATAGACTTT	1258
Db	743	ThrAsnThrThrLeuSerLeuLeuProGluProLysPro-----	755
QY	1259	GTCTGTACCTGCCAAGGAGCATTCACCGAGGTC---TGTACCATCATTTCTGACCCC	1315
Db	756	-----ProThrGluValGlnSerProLeuMetAlaAspPro	767
QY	1316	ACCTGCCAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGTGTCTGTG	1375
Db	768	ThrCys-----ValGlyProAlaValSerPheSerSer-----	778
QY	1376	CTGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGG	1435
Db	779	LeuTyrProArgAspPhePheProAsnThrAlaSerPheLeuThrArgLeuArgLeuSer	798
QY	1436	GATGACACAAGCCTGGCTCTCACGAGCACC-----CTGATTCTGTTCCTGACAGA	1486
Db	799	ProProThrProLeuProMetProLysAsnAsnCysLeuLeuThrAlaValAlaProSer	818
QY	1487	GACCAGCCTCGCCTTTAAGG-----	1507
Db	819	LeuHisIleAsnProHisArgLeuTrpThrSerLeuGlnGluValLeuProAspSerLeu	838
QY	1508	ATGGCAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTGTCACTGTGATCTCC	1567
Db	839	LeuSerAsnSerGluIleAspSerValGlyMetSerThrAspLeuLeuThrAlaLeuSer	858

US-10-039-272-1 (1-2661) x KGHUGH (1-525)

```

QY 1028 CCGCCACCGCCACCA-----CCACCAGACCTTCAAAACCC 1063
    |||||||
Db 297 ProProProProAspGluArgAspHisSerHisGlyProProLeuProGlnGlyPro 316
    |||||||
QY 1064 ACCCCTTCTTAGGACCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCTGTATGAA 1123
    ||| ||| |||:::
Db 317 ProProLeuLeu-----ProMetSerCysSer----- 325
    ||| ||| |||

QY 1124 AACTGCCAGATTAAACAGATATGGCCA----- 1149
    :::|||||
Db 326 SerCysGlnHisAlaThrPheGlyThrAsnGlyAlaGlnArgHisSerHisAsnAsn 345
    ||| ||| ||| |||:::

QY 1150 -----CTTCAAGCCAC--CATCACAATTGTAGAGGGAATCTTAGAGGTTAACAT 1197
    ||| ||| ||| |||:::
Db 346 SerSerAspLeuHisProHisLysHisHisSerHisGluGlnHisProHisGly--His 364
    ||| ||| ||| ||| |||

QY 1198 CATCCAGATGACAGACGTCCTGATGCCGGTGCCATGGCCCTGAAAGCTCCCTAATAGACTT 1257
    |||||||
Db 365 HisProHisAlaHisHisProHisGlu----- 373
    ||| ||| |||

QY 1258 TGTCTGTGACCTGCCAAGGGAGCATTCACACGGAGGTCTGTACCATCATTTCTGACCCAC 1317
    ||| |||::: ||| ||| |||
Db 374 ---HisAspThrHisArgGlnHisProHisGly-----HisHis-----ProHis 387
    ||| ||| ||| ||| |||

QY 1318 CTGCGAGATCACCCGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCT 1377
    ||| ||| ||| |||
Db 388 -----GlyHisHisProHisGlyHisHisProHisGly----- 398
    ||| ||| ||| |||

QY 1378 GACTGTGAGACGAACCTTCAATGGTCTGCGGACGTACTGTGTGAACCTCACCTGGGGGA 1437
    ||| ||| ||| ||| |||
Db 399 -----HisHisProHisGly 403
    ||| ||| ||| |||

QY 1438 TGACACAAGCCTGGCTCTCACGAGACCCCTGATTTCTGTCTGTGACAGAGACCCAGCCTC 1497
    ||| ||| |||:::
Db 404 ---HisHisProHisCysHisAsp----- 410
    ||| ||| |||

QY 1498 GCCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGCTTGCCATATTGTTCAC 1557
    |||::: ||| ||| ||| |||
Db 411 ---PheGlnAspTyrGlyProCysAspProProHisAsnGlnGlyHisCysCysHis 429
    |||::: ||| ||| ||| |||

RESULT 15
JQ0533
genome polyprotein - Kennedy yellow mosaic virus (strain Jarvis Bay)
N;Alternate names: RNA nucleotidyltransferase (RNA-directed); RNA replicase
N;Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C;Species: Kennedy yellow mosaic virus
C;Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 19-Jan-2001
C;Accession: JQ0533
R;Ding, S.; Keese, P.; Gibbs, A.
J. Gen. Virol. 71, 925-931, 1990
A;Title: The nucleotide sequence of the genomic RNA of Kennedy yellow mosaic tymovirus
A;Reference number: JQ0532; MUID:90218040; PMID:2324710
A;Accession: JQ0533
A;Molecule type: Genomic RNA
A;Residues: 1-1874 <DIN>
A;Cross-references: GB:D00637; NID:g221969; PIDN:BAA00532.1; PID:g221971
C;Superfamily: eggplant mosaic virus RNA-directed RNA polymerase
C;Keywords: ATP; nucleotide binding; nucleotidyltransferase; P-loop; RNA biosynthesis; F;
F;1002-1009/Region: nucleotide-binding motif A (P-loop)
F;1064-1069/Region: nucleotide-binding motif B
F;1008/Binding site: ATP (Lys) #status predicted

Alignment Scores:
Pred. No.: 0.652 Length: 1874
Score: 110.50 Matches: 77
Percent Similarity: 31.22% Conservative: 36
Best Local Similarity: 21.27% Mismatches: 112
Query Match: 2.28% Indels: 137
DB: 1 Gaps: 15

```

Qy 1568 CTCTTG 1573
|||
Db 859 HisLeu 860

Search completed: September 10, 2004, 15:51:13
Job time : 108.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2004, 01:04:08 ; Search time 10036 Seconds
(without alignments)
7917.822 Million cell updates/sec

Title: US-10-039-272-1
Perfect score: 2661
Sequence: 1 cggcacgagggccagagga.....aaaaaaaaaaaaaaaaaaaaa 2661

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2550	95.8	2636	11 BC025297	BC025297 Homo sapi
2	1049.8	39.5	3615	11 AK044764	AK044764 Mus muscu
3	1046.6	39.3	2265	11 AK076347	AK076347 Mus muscu
4	1025	38.5	1201	13 BX397314	BX397314 BX397314

C	5	1014.6	38.1	1052	13	BX379978	BX379978
	6	943.8	35.5	2431	11	AK079220	Mus muscu
C	7	937	35.2	1003	13	BX423077	BX423077
C	8	926.6	34.8	1015	13	BX458448	BX458448
	9	918.8	34.5	1139	13	BX406949	BX406949
	10	914	34.3	1054	13	BX379979	BX379979
	11	910.6	34.2	1201	13	BX336884	BX336884
	12	906.4	34.1	1201	13	BX364871	BX364871
	13	885.6	33.3	1012	13	BX458449	BX458449
C	14	868	32.6	1138	13	BX406948	BX406948
	15	867.4	32.6	895	14	CF552020	AGENCOURT
	16	864.4	32.5	1002	13	BX423078	BX423078
C	17	862.2	32.4	927	9	AL542811	AL542811
	18	854	32.1	1096	13	BX381217	BX381217
C	19	852	32.0	1201	13	BX364870	BX364870
C	20	835.4	31.4	1048	9	AL575920	AL575920
C	21	831.6	31.3	972	13	BX396828	BX396828
	22	823	30.9	923	9	AL542812	AL542812
	23	816.6	30.7	891	13	BX452668	BX452668
	24	815.8	30.7	911	12	BI521316	603081887
	25	799.8	30.1	972	13	BX396829	BX396829
	26	796.8	29.9	858	12	BG742951	BG742951
	27	788	29.6	1107	12	BM550298	AGENCOURT
	28	779	29.3	1089	12	BM547872	AGENCOURT
	29	776	29.2	835	12	BG576651	602597538
	30	775	29.1	808	12	BG742272	602631402
	31	775	29.1	962	12	BG676576	602623049
	32	758.2	28.5	810	9	AU139997	AU139997
C	33	750.4	28.2	1201	9	AL564958	AL564958
	34	745	28.0	754	12	BG211992	RST31703
	35	743.8	28.0	892	13	BUI57659	AGENCOURT
C	36	742.4	27.9	878	13	BX437190	BX437190
	37	737.8	27.7	855	13	BUI78305	AGENCOURT
	38	733.8	27.6	807	12	BG696053	602658007
	39	731.6	27.5	905	14	CD109017	AGENCOURT
	40	728.2	27.4	857	13	BUI50501	AGENCOURT
	41	722.2	27.1	808	14	CD108927	AGENCOURT
C	42	715.2	26.9	941	9	AL550419	AL550419
	43	714.8	26.9	745	12	BG742920	602632010
	44	714.8	26.9	793	13	BUI95030	AGENCOURT
C	45	714.8	26.9	891	13	BX412287	BX412287

ALIGNMENTS

RESULT 1
BC025297
LOCUS BC025297
DEFINITION Homo sapiens, glycoprotein (transmembrane) nmb, clone
IMAGE:4877773, mRNA.
ACCESSION BC025297
VERSION BC025297.1 GI:19264140
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

BC025297 2636 bp mRNA linear HTC 08-MAR-2002
Homo sapiens, glycoprotein (transmembrane) nmb, clone

REFERENCE 1 (bases 1 to 2636)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Brin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 42 Row: i Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4505404
This clone has the following problem: frame shifted.

FEATURES
source
location/Qualifiers
1. .2636
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:10457"
/db_xref="taxon:9606"
/clone="IMAGE:4877773"
/tissue_type="Skin, melanotic melanoma, high MDR."
/clone_lib="NIH MGC 49"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

ORIGIN
Query Match 95.8%; Score 2550; DB 11; Length 2636;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 2595; Conservative 0; Mismatches 20; Indels 3; Gaps 3;

QY 28 AACCTTGGTGGCTCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGGAT 87
Db 20 ACCTTGAGTGCCTCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGGAT 79
QY 88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCCGCCAAACGATTTTCATGATGCTGG 147
Db 80 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCCGCCAAACGATTTTCATGATGCTGG 139
QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTAATGGCTGGTCTTCTG 207
Db 140 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTAATGGCTGGTCTTCTG 199
QY 208 ATGAAATGACTGGAATGAAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 267
Db 200 ATGAAATGACTGGAATGAAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGAA 259
QY 268 AAAACTCCTGGAAGGAGCGGTGTGCAGGCGGTCTGACCACTGACTCACCAGCCCTCG 327
Db 260 AAAACTCCTGGAAGGAGCGGTGTGCAGGCGGTCTGACCACTGACTCACCAGCCCTCG 319
QY 328 TGGGCTCAAATATAACATTTGCGGTGAACCTGATATCCCTAGATGCCAAAAGGAAGATG 387
Db 320 TGGGCTCAAATATAACATTTGCGGTGAACCTGATATCCCTAGATGCCAAAAGGAAGATG 379
QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 447
Db 380 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 439
QY 448 CATATGTTTACAACCTGGACAGCATGGTTCAGAGGACAGTGCAGGGGAAAAATGGCACCGGCC 507
Db 440 CGTATGTTTACAACCTGGACAGCATGGTTCAGAGGACAGTGCAGGGGAAAAATGGCACCGGCC 499
QY 508 AAAGCCATCATAAACGTCTTCCCTGATGGGAAACCTTTTCTCCTCACCACCCCGGATGGAGAA 567
Db 500 AAAGCCATCATAAACGTCTTCCCTGATGGGAAACCTTTTCTCCTCACCACCCCGGATGGAGAA 559
QY 568 GATGGAATTTTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT 627
Db 560 GATGGAATTTTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT 619

QY 628 GTTCAGTGAGAGTTTCTGTGAACACAGCCAATGTGACACTTGGGCCTCAACTCATGGAG 687
Db 620 GTTCAGTGAGAGTTTCTGTGAACACAGCCAATGTGACACTTGGGCCTCAACTCATGGAG 679
QY 688 TGACTGTCTACAGAAAGACATGGACGGGCATATGTTCCATCGCACAAAGTGAAGATGTGT 747
Db 680 TGACTGTCTACAGAAAGACATGGACGGGCATATGTTCCATCGCACAAAGTGAAGATGTGT 739
QY 748 ACGTGGTAAACAGATCAGATTCTCTGTGTTTGTGACTATGTTCCAGAAGAACGATCGAAAT 807
Db 740 ACGTGGTAAACAGATCAGATTCTCTGTGTTTGTGACTATGTTCCAGAAGAACGATCGAAAT 799
QY 808 CATCCGACGAAACCTTCC-CAAGATCTCCCCATTTATGTTGATGTCTCTGATTCAATGATC 866
Db 800 CATCCGACGAAACCTTCCCTCAAGATCTCCCCATTTATGTTGATGTCTCTGATTCAATGATC 859
QY 867 CTAGCCACTTCTCAATTATTTCTACCATTAACAAAGTGGAGCTTCGGGGATAATACTG 926
Db 860 CTAGCCACTTCTCAATTATTTCTACCATTAACAAAGTGGAGCTTCGGGGATAATACTG 919
QY 927 GCCTGTTGTTTCCACCAATCACTGTGAATCACACGATATGCTCAATGGAACCTTCA 986
Db 920 GCCTGTTGTTTCCACCAATCACTGTGAATCACACGATATGCTCAATGGAACCTTCA 979
QY 987 GCCTTAACTCACTGTGAAAGCTGCAGCACCAAGGACCTTGTCCGCCACCGCCACCAC 1046
Db 980 GCCTTAACTCACTGTGAAAGCTGCAGCACCAAGGACCTTGTCCGCCACCGCCACCAC 1039
QY 1047 CCAGACCTTCAAAAACCCCTTCTTTAGGACCTGCTGGTGACAAACCCCTGGAGCTGA 1106
Db 1040 CCAGACCTTCAAAAACCCCTTCTTTAGGACCTGCTGGTGACAAACCCCTGGAGCTGA 1099
QY 1107 GTAGGATTCCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCA 1166
Db 1100 GTAGGATTCCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCA 1159
QY 1167 CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGG 1226
Db 1160 CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGG 1219
QY 1227 TGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGCCAAGGGAGCATTTCCCA 1286
Db 1220 TGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGCCAAGGGAGCATTTCCCA 1279
QY 1287 CGGAGGTCTGTACCATCATTTCTGACCCCTGCGGAGATCACCCAGAACACAGTCTGCA 1346
Db 1280 CGGAGGTCTGTACCATCATTTCTGACCCCTGCGGAGATCACCCAGAACACAGTCTGCA 1339
QY 1347 GCCCTGTGGATGTGGATGAGATGTGCTGTGACTGTGAGACGAACTTCAATGGGTCTG 1406
Db 1340 GCCCTGTGGATGTGGATGAGATGTGCTGTGACTGTGAGACGAACTTCAATGGGTCTG 1399
QY 1407 GGACGTACTGTGTGAACCTCACCTGGGGGATGACAAAGCCTGGCTCTCACGAGCACCC 1466
Db 1400 GGACGTACTGTGTGAACCTCACCTGGGGGATGACAAAGCCTGGCTCTCACGAGCACCC 1459
QY 1467 TGATTTCTGTTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAACAGTGCCTGA 1526
Db 1460 TGATTTCTGTTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAACAGTGCCTGA 1519
QY 1527 TCTCCGTTGGCTGTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAC 1586
Db 1520 TCTCCGTTGGCTGTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAC 1579
QY 1587 ACAAGGAATACAAACCCCAATAGAAAAATAGTCTGGGAAATGTTGGTTCAGAAAGGCTGA 1646
Db 1580 ACAAGGAATACAAACCCCAATAGAAAAATAGTCTGGGAAATGTTGGTTCAGAAAGGCTGA 1639
QY 1647 GTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCGGGAAACCCAGGAAAGGATCCGC 1706
Db 1640 GTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCGGGAAACCCAGGAAAGGATCCGC 1699
QY 1707 TACTCAAAAACCAAGAAATTTAAAGGAGTTTCTTAAATTTTCGACCTGTTTCTGAAGCTCA 1766

Db	1700	TACTCAAAAACCAAGAAATTTAAAGGAGTTTCTTAAATTTTCGACCTGTGTTTCTGAAGCTCA	1759
QY	1767	CTTTTCAGTGCCTATTGATGTGAGATGTCTGGAGTGGCTATTAACTTTTTTTCCTAAAG	1826
Db	1760	CTTTTCAGTGCCTATTGATGTGAGATGTCTGGAGTGGCTATTAACTTTTTTTCCTAAAG	1819
QY	1827	ATTATTGTTAAATAGATATTGTGGTTTGGGAAAGTTGAAATTTTATAGGTTAAATGTCA	1886
Db	1820	ATTATTGTTAAATAGATATTGTGGTTTGGGAAAGTTGAAATTTTATAGGTTAAATGTCA	1879
QY	1887	TTTTAGAGATGGGAGAGGGAATTATACTGTCAGGCAGCTTCAGCCATGTTGTGAAACTGAT	1946
Db	1880	TTTTAGAGATGGGAGAGGGAATTATACTGTCAGGCAGCTTCAGCCATGTTGTGAAACTGAT	1939
QY	1947	AAAAGCAACTTAGCAAGGCTTCTTTTCATTATTTTATGTTTCACTTATAAAGTCTTAG	2006
Db	1940	AAAAGCAACTTAGCAAGGCTTCTTTTCATTATTTTATGTTTCACTTATAAAGTCTTAG	1999
QY	2007	GTAAGTGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAACTACTATTGATTA	2066
Db	2000	GTAAGTGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAACTACTATTGATTA	2059
QY	2067	GAGCCTAACCCAGGTTAACTGCAAGAAGAGCGGGATACTTTTCAGCTTTCCATGTAAGT	2126
Db	2060	GAGCCTAACCCAGGTTAACTGCAAGAAGAGCGGGATACTTTTCAGCTTTCCATGTAAGT	2119
QY	2127	TATGCATAAAGCCAAATGAGTCCAGTTTCTAAGATCATGTTTCCAAAGCTAACTGAATCCCA	2186
Db	2120	TATGCATAAAGCCAAATGAGTCCAGTTTCTAAGATCATGTTTCCAAAGCTAACTGAATCCCA	2179
QY	2187	CTTCAATACACACTCATGAACCTCCTGATGGAACAATAACAGGCCCAAGCCTGGGTATGA	2246
Db	2180	CTTCAATACACACTCATGAACCTCCTGATGGAACAATAACAGGCCCAAGCCTGGGTATGA	2239
QY	2247	TGTGCACACTTGCTAGACTCAGAAAAAATACTACTCTCATATAAATGGTGGGAGTATTTTG	2306
Db	2240	TGTGCACACTTGCTAGACTCAGAAAAAATACTACTCTCATATAAATGGTGGGAGTATTTTG	2299
QY	2307	GTGACAAACCTACTTTCCTTGGCTGAGTGAAGGAATCATATATTCATATTTATTCCTCA	2366
Db	2300	GTGACAAACCTACTTTCCTTGGCTGAGTGAAGGAATCATATATTCATATTTATTCCTCA	2359
QY	2367	TGGACATTTAGTTAGTGCCTTTTATATATACCAGGCATGATGCTGAGTGACACTCTTGTGTA	2426
Db	2360	TGGACATTTAGTTAGTGCCTTTTATATATACCAGGCATGATGCTGAGTGACACTCTTGTGTA	2419
QY	2427	TATTTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAAATATTAAGACTTTCC	2486
Db	2420	TATTTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATC-ATATATTAAGACTTTCC	2478
QY	2487	AAAAATTTGGTCCTGGTTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTG	2546
Db	2479	AAAGATGAGTCCCTGGTTTTTTCATGGCAACTTGATCAGTAAGGATTTCACTCTGTTTG	2538
QY	2547	GAACTAAACCATTTTACTATATGTTTAGACAAGACATTTTTTTTTTTCCTTCCTGAAAAA	2606
Db	2539	TAACTAAACCATCTACTATATGTTTAGACATGACATTTCTTTTCTCTCCTTCCTGAAAAA	2598
QY	2607	-AAAATGCGGAAGAGACAAAAAATAAAAAAATAAAAAA 2643	
Db	2599	TAAAGTGTGGAAGAGACAAAAAATAAAAAAATAAAAAA 2636	
RESULT 2			
AK044764			
LOCUS	AK044764	3615 bp	mRNA
DEFINITION	Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone:A930040L12 product:glycoprotein (transmembrane) nmb, full insert sequence.	linear	HTC 20-SEP-2003
ACCESSION	AK044764		
VERSION	AK044764.1	GI:26336782	
KEYWORDS	HTC; CAP trapper.		

SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS	1 Carninci,P. and Hayashizaki,Y.
TITLE	High-efficiency full-length cDNA cloning
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)
MEDLINE	99279253
PUBMED	10349636
AUTHORS	2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE	20499374
PUBMED	11042159
AUTHORS	3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE	20530913
PUBMED	11076861
AUTHORS	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
JOURNAL	Nature 409, 685-690 (2001)
REFERENCE	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
AUTHORS	Analysis of the mouse transcriptome based on functional annotation Of 60,770 full-length cDNAs
TITLE	Nature 420, 563-573 (2002)
JOURNAL	6 (bases 1 to 3615)
REFERENCE	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
AUTHORS	Direct Submission
TITLE	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
JOURNAL	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
COMMENT	Retina RNA was provided by Dr. Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged. Please visit our web site for further details. URL:http://genome.gsc.riken.go.jp/

URL:http://fantom.gsc.riken.go.jp/.

FEATURES

Location/Qualifiers

1. .3615

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM DB:A930040L12"

/db_xref="MGI:2410100"

/db_xref="taxon:10090"

/clone="A930040L12"

/tissue_type="retina"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

100. .1824

/note="unnamed protein product; glycoprotein (transmembrane) nmb (MGD|MGI:193476S, GB|NM_053110, evidence: BLASTN, 99%, match=2272)

putative"

/codon_start=1

/protein_id="BAC32074.1"

/db_xref="GI:26336783"

/translation="MESLGVGLFLLLAAGLPLOAAKRFDRVLDGHEQYPNHMREHNOLRGWSSDENWEHLYPVWRRGDRWKDSWEGGRVQAVLTSDSPALVGSNITFVNLFVPRCKEDANGNIVYEKNCRNDLGLTSDLHVNTAGADGDWEDGTSRSQHLRFPDRRPFPRPHGWKWSFYVFTLQGYFQKLGRCSARVSINTVNLTAGPQVMEVTVFRRYGRAYIPISKVDYVITDQIPFVFTMSQKNDRLNLSDEIFLRDLPIFVFDVLIHDPSSHFLNDSAISYKNWFGDNTGLFVSNHHTLNHTYVLNGTFLNLTQVTAVPGPCPPSPSTPPPPSTPSPSPPLTLSTPSPSLMPTGYKSMELSDISNENCRINRYGYFRATITIVIGILEVSIQIADVPMPTPPANSLMDFTVCKGATPEACTIISDPTCQIAQNRVCSPVAVDGLCLSVRRAPNGSGTYCVNFTLGDASIALTSLISIPGKDPDSPMRAVNGVLISIGCLAVLTMVTILLYKKHKAYKPIGNCPRNTVEGKGLSVLLSHAKAPFFRQDQEKDPLLQDKPRTL"

polyA_signal

3600. .3605

/note="putative"

polyA_site

3615

/note="putative"

ORIGIN

Query Match 39.5%; Score 1049.8; DB 11; Length 3615;

Best Local Similarity 74.0%; Pred. No. 4.5e-160;

Matches 1409; Conservative 0; Mismatches 442; Indels 54; Gaps 4;

QY 30 CCTTGGTCCGTCCGTGAGAAATTCAGCATGGAAATGTCTCTACTATTTCCTGGGATTT 89

Db 70 CCTGACTGGTTGCAGGCGCTCGGAGTCAGCATGGAAAGTCTCTGCGGGGTCTCTGGGATTT 129

QY 90 CTGCTCCTGGCTGCAAGATTGCCACTTGTATGCCGCCAAACGATTTTCATGATGTGCTGGGC 149

Db 130 CTGCTGCTGGCTGCAGGACTGCCTCTCCAGGCTGCCAAGCGATTTTCGTGATGTGCTGGGC 189

QY 150 AATGAAAGACCTTCTGCTTACATGAGGGAGCAACAATCAATATAATGGCTGGTCTTCTGAT 209

Db 190 CATGAACAGTATCCCAATCATGAGAGAGCAACAACCAATTACGTGGCTGGTCTTCGGAT 249

QY 210 GAAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGGAAA 269

Db 250 GAAAATGAATGGGATGAACACCTGTATCCAGTGTGGAGGAGGGAGACGGCAGGTGGAAG 309

QY 270 AACTCTGGAAGGAGGCGCGTGTGCAGGCGGTCTGACCACTGACTCACCAGCCCTCGTG 329

Db 310 GACTCTGGGAGGAGGCGCGTGTGCAGGCGAGTCTGACCACTGACTCACCAGCTCTGGTG 369

QY 330 GGCTCAAAATATAACATTTGGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATGCC 389

Db 370 GGTTCCAATATCACCTTTGTGGTGAACCTGGTGTCCCAAGATGCCAGAAGGAAGATGCT 429

QY 390 AATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATCCA 449

Db 430 AATGGCAATATCGTCTATGAGAAGAACTGCAGGAATGATTTGGGACTGACCTCTGACCTG 489

QY 450 TATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGACGGGGAAAAATGGCACCGGCCAA 509

Db 490 CATGCTCTACAACTGGACTGCAGGGGCAGATGATGGTGAATGGTGAAGATGGCACCGCCA 549

QY	510	ACCCATCATAAACGCTTTCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAGA 569
Db	550	ACCCAGCATCTCAGGTTCCCGACAGGAGGCCCTTCCCTCGCCCCCATGGATGAAGAA 609
QY	570	TGGAATTTTCATCTACGTTCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGATGT 629
Db	610	TGGAGCTTTGTCTACGTTCTTACACACTTGGCCAGTATTTCCAAAACTGGTGGTGGTGT 669
QY	630	TCAGTGAGAGTTTCTGTGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAGTG 689
Db	670	TCAGCACGGGTTTCTATAAACACAGTCAACTTGACAGCTGGCCCTCAGGTCTAGGAAGTG 729
QY	690	ACTGTCTACAGAAAGACATGGACGGGCATATGTTCCCATCGACAAGTGAAGATGTGTAC 749
Db	730	ACTGTCTTTCGAAGATACGGCCGGGCATACATTCCCATCTCGAAGGTGAAGATGTGTAT 789
QY	750	GTGTTAAACAGATCAGATTCTCTGTGTTGTGACTATGTTCCAGAGAACGATCGAAATTCA 809
Db	790	GTGATAACAGATCAGATCCCTGTATTCTGTGACCATGTCCAGAAATGACAGGAACCTTG 849
QY	810	TCCGACGAAACCTTCC-CAAAGATCTCCCCCATATTATGTTGATGTCTGATTCATGATCCT 868
Db	850	TCTGATGAGATCTTCTTCAGAGACCTCCCCCATCGTCTTCGATGTCTCTATTGATCCTCC 909
QY	869	AGCCACTTCTCTCAATATTATTCTACCAATTAACAAAGTGGAGCTTCGGGATTAATACTGGC 928
Db	910	AGCCACTTCTCTCAACGACTCTGCCATTCTCTACAAGTGGAACCTTTGGGACAACTGGC 969
QY	929	CTGTTTGTGTTCCACCAATCATACTGTGAATCACACGATATGTGCTCAATGGAACCTTCAGC 988
Db	970	CTGTTTGTCTCCAAACAATCACACTTTTGATCATCACTTATGTCTCAATGGAACCTTCAAC 1029
QY	989	CTTAACCTCACTGTGAAAGCTGCAGCACAGGACCTTGTCCGCCACCGC----- 1037
Db	1030	CTTAACCTCACCGTGCAAACTGCAGTGCCCGGGCCATGCCCTCCCCCTTCGCTTCGACT 1089
QY	1038	-----CACCACCAACCCAGACCTTCAAAA 1060
Db	1090	CCGCTCCACCTTCAACTCGCCCTCACCTCCGCTCCGCTCACCTTGCCTTATCAACA 1149
QY	1061	CCACCCCTTCTTTAGGACCTGTGGTGACAACCCCTGGAGCTGAGTAGGATTCCTGAT 1120
Db	1150	CCTAGCCCTCTTTAATGCCTACTGTTTACAATCCATGGAGCTGAGTGACATTTCCAAT 1209
QY	1121	GAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCACAAATTTAGAGGGA 1180
Db	1210	GAAAACTGCCGAATAAACAGATATGGCTACTTTCAGAGCCACCATCACAAATTTAGAGGGG 1269
QY	1181	ATCTTAGAGGTTAAACATCATCCAGATGACAGAGCTCTGATGCCGTTGCCATGGCTGAA 1240
Db	1270	ATCCTGGAAGTCAGCATCATGAGATAGAGATGTCCCCATGCCACACCCGACGCTGCC 1329
QY	1241	AGCTCCCTAATAGACTTTGTGTCGACCTGCCAAGGAGGACATTCACACGGAGGTCTGTACC 1300
Db	1330	AACTCCCTGATGGACTTCACTGTGACCTGCAAGGGGCGCCACCCCATGGAAGCCTGTACG 1389
QY	1301	ATCATTTCTGACCCCACTTCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTG 1360
Db	1390	ATCATCTCCGACCCCACTTCGAGATCGCCAGAACCCGGTCTGCAGCCCTGTGGCTGTG 1449
QY	1361	GATGAGATGTCTGTGCTGCTGTGAGACGAACTTCAATGGGTCTGGACGTAATGTGTG 1420
Db	1450	GAITGGCTGTGCTGTGCTGTGAGAGAGCCTTCAATGGGTCTGGCACCTACTGTGTG 1509
QY	1421	AACCTCACCTGGGGATGACAAAGCCTGGCTCTCACAGACACCCCTGATTTCTGTTCCT 1480
Db	1510	AATTTCACTCTGGGAGATGATGCAAGCCTGGCCCTCACCCAGACCCCTGATCTCTATCCCT 1569
QY	1481	GACAGAGACCCAGCCTCGCCCTTAAAGGATGGCAACACAGTGCCTGATCTCCGTTGGCTGC 1540
Db	1570	GGCAAGACCCAGACTCCCCTATGAGAGCAGTGAATGGTGTCTCTGATCTCCATTGGCTGC 1629

QY 1541 TTGGCCATATTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAAGGAATACAAC 1600
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1630 CTGGCTGTGCTTGTCAACCATGGTTACCATCTTGCTGTACAAAAACACAAGCGGTACAAG 1689

QY 1601 CCAATAGAAAATAGTCTCTGGGAATGTGGTCAGAAAGCAAAGCGCTAGTGTCTTCTCAAC 1660
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1690 CCAATAGAAAATGCCCCAGGAACACGCTCGAGGCAAAGCGCTAGTGTCTCTCCTCAGC 1749

QY 1661 CGTGCAAAAGCCGTGTTCTTCCCGGAAACACGAGAAAGGATCCGCTACTCAAAAACCAA 1720
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1750 CACGCGAAAGCCCGTCTTCCGAGGAGACCAGGAGAGGATCCATTGCTCCAGGACAAG 1809

QY 1721 GAATTTAAAGGAGTTTCTTAAATTCGACCTTGTTTCTGAAGCTCACTTTTCAGTGCCAT 1780
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1810 CCAAGGACACTCTAAGTCTTTGGCGCTTCCCTCTGACCAGGAACCCACTCTTCTGTG-CAT 1868

QY 1781 TGATGTGAGATGTGCTGGA-----GTGGCTATTAACTTTTTCCTAAAGATTATTGTTA 1836
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1869 GTATGTGAGCTGTGCAGAAAGTATGTGGCTGGGAACTGTTGTTCTTAAGGATTATTGTAA 1928

QY 1837 AATAGATATTGTGTTTGGGGAAGTTGAATTTTATTATAGGTTAAA 1881
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1929 AATGTATATCGTGGCTTAGGAGTGTGGTTAAATAGCAATTTAGA 1973

RESULT 3
AK076347
LOCUS
DEFINITION AK076347 2265 bp mRNA linear HTC 18-SEP-2003
Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched
library, clone:4732452017 product:glycoprotein (transmembrane) nmb,
full insert sequence.
AK076347
ACCESSION AK076347.1 GI:26345309
VERSION HTC; CAP trapper.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
REFERENCE 2
AUTHORS Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20499374
PUBMED 11042159
REFERENCE 3
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL Genome Res. 10 (11), 1757-1771 (2000)
MEDLINE 20530913
PUBMED 11076861
REFERENCE 4
AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409, 685-690 (2001)
REFERENCE 5
AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.

TITLE
JOURNAL
REFERENCE
AUTHORS

Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2265)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Haragaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.

TITLE
JOURNAL

Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

COMMENT

FEATURES
source

Location/Qualifiers
1. .2265
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:4732452017"
/db_xref="MGI:2390753"
/db_xref="taxon:10090"
/clone="4732452017"
/tissue_type="skin"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 days neonate"
75. .1799
/note="unnamed protein product; glycoprotein
(transmembrane) nmb (MGD|MGI:1934765, GB|NM_053110,
evidence: BLASTN, 99%, match=2272)
putative"
/codon_start=1
/protein_id="BAC36306.1"
/db_xref="GI:26345310"
/translation="MESLCCVLGFLLLAAGLPLOAKRFRDVLGHEQYPDHMRHNQL
RGWSSDENEDHLYPVWRGDGRWKDSWEGRVQAVLTSDPALVGSNITFVVNLVF
PRCKEDANGNI VYEKNCRNDLGLTSDLHVYNNWTAGDDGDWEDGTSRSQHLRFPDRR
PEPRPHGKKWSFVYVFHTLQGYFQKLGRCRSARVSINTVNLTAGPQVMEVTVFRYGR
AYIPISKVDYVITDQIPVFVTMSQXNDRNLSDEIFLRLDPIVEDVLIHDPSEHFLND
SAISYKWNFGDNTGLFVSNHNLNHTYVLNGTFNLNLTVQTAVPGPCPPSPSTPPP
STPPSPSPPLTSPSLMPTGYKSMELSDISNENCRINRYGYFRATITIVGIL
EVSIMQIADVPMPTQDANSLMDFVTCKGATPMEACTIISDPTCQIAQNRVCSPVAV
DGLCLLSVRRAFNGSGTCVNFITLGDASLALTSILISIPGKDPDSPLRVNGVLISI
GCLAVLVTMTVITLLYKKHKAYKPIGNCPRNTVKGGLSVLLSHAKAPFRFGDQEKDPL
LQDKPRTL"
2250. .2255
/note="putative"
2265
/note="putative"

CDS

polyA_signal

2250. .2255

/note="putative"

polyA_site

2265

/note="putative"

ORIGIN

Query Match 39.3%; Score 1046.6; DB 11; Length 2265;
Best Local Similarity 73.9%; Pred. No. 1.9e-159;
Matches 1407; Conservative 0; Mismatches 444; Indels 54; Gaps 4;

QY 30 CTTGGTGCCTCGCTCGGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGGATTT 89
Db 45 CTTGACTGGTTGCAGCGCTCGGAGTCAGCATGGAAGTCTCTCGGGGTCCTGGGATTT 104
QY 90 CTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTCATGATGTCTGGGC 149
Db 105 CTGCTGCTGGCTGCAGGACTGCCTCTCCAGGCTGCCAAGCGAATTCGTGATGTCTGGGC 164
QY 150 AATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGGTCTTCGAT 209
Db 165 CATGAACAGTATCCCGATCACATGAGAGAGCACAAACCAATTACGTGGCTGGTCTTCGGAT 224
QY 210 GAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGGAAA 269
Db 225 GAAATGAATGGATGAACACCTGTATCCAGTGTGAGGAGGGGAGACGGCAGGTGGAA 284
QY 270 AACTCCTGGAAGGGAGCGCTGTGACGGCGGTCTCTGACAGTGAATCACCAGCCCTCGTG 329
Db 285 GACTCCTGGGAAGGAGCGCTGTGACGGCAGTCTCTGACAGTGAATCACCAGCTCTGGTG 344
QY 330 GGCTCAATATACATTTGCGGTGAACCTGATATTTCCCTAGATGCAAAAGGAAGATGCC 389
Db 345 GGTTCCAATATACATTTTGTGGTGAACCTGGTGTCTCCAGATGCCAGAAGAGATGCT 404
QY 390 AATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGGTGTATCTGCTGATCCA 449
Db 405 AATGGCAATATCGTCTATGAGAAGAACTGCAGAAATGATTTGGGACTGACATCTGACCTG 464
QY 450 TATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGACGGGGAATAATGGCACCGGCAA 509
Db 465 CATGTCTACAACTGGACTGCAGGGGCAGATGATGGTGAATGGGAGATGGCACCGCGA 524
QY 510 AGCCATCATACCTCTTCCCTGATGGGAACCTTTTCCCTCACACCCCGGATGGAGAAGA 569
Db 525 AGCCAGCATCTCAGGTTCCCGACAGGAGGCCCTTCCCTCGCCCCCATGGATGGAAGAAA 584
QY 570 TGGAAATTTCACTACGTCTTCCACACACTTGTGTCAGTATTTCCAGAAATTTGGACGATGT 629
Db 585 TGGAGCTTTGTCTACGTCTTTCACACACTTGGCCAGTATTTCCAAAACCTGGTGGTGT 644
QY 630 TCAGTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCCCTCAACTCATGGAAGTG 689
Db 645 TCAGCACGGGTTTCTATAAACACACAGTCAACTTGACAGCTGGCCCTCAGGTGATGGAAGTG 704
QY 690 ACTGTCTACAGAAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAACATGTGTAC 749
Db 705 ACTGTCTTTTCGAAGATACGGCCGGGCATACATTCCTCATCTCGAAGGTGAAGATGTGTAT 764
QY 750 GTGGTAACAGATCAGATTCTCTGTGTTGTGACTATGTTCCCATCGCACAAAGTGAACATTC 809
Db 765 GTGATAACAGATCAGATCCCTGTATTCTGTGACCATGTGCCAGAAGAAATGACAGAACTTG 824
QY 810 TCCGACGAAACCTTCC-CAAAGATCTCCCCATATGTTGATGTCTGATTCATGATCCT 868
Db 825 TCTGATGAGATCTTCTCTCAGAGACCTCCCCCATCGTCTTCGATGTCCTCATTCATGATCCC 884
QY 869 AGCCACTTCTCAATATTCTACCATTAACATAAGTGAGCTTCGGGGATAATACCTGGC 928
Db 885 AGCCACTTCTCAACGACTCTGCCATTTCTTACAGTGAACCTTTGGGGACAAACACTGGC 944
QY 929 CTGTTTGTCTCCACCAATCATACTGTGAATCACACGATGTGCTCAATGGAACCTTCAGC 988
Db 945 CTGTTTGTCTCCAACAATCACACTTTGAATCACACTTATGTGCTCAATGGAACCTTCAAC 1004
QY 989 CTTAACCTCACTGTGAAAGCTGCAGCACAGGACCTTGTTCGCCACCGC----- 1037
Db 1005 CTTAACCTCACTGTGCAAACTGAGTGGCCGGGCCATGCCCTCCCCCTTCGCCCTTCGACT 1064
QY 1038 -----CACCAACCCACAGACCTTCAAAA 1060
Db 1065 CCGCCTCCACCTTCAACTCCGCCCTCACCTCCGCCCTCACCTCTGCCCAATATCAACA 1124
QY 1061 CCCACCCCTTCTTTAGGACCTGTCTGGTGACAACCCCTCGGAGCTGAGTAGGATTCCTGAT 1120

Db 1125 CCTAGCCCCCTCTTTAATGCCTACTGGTTACAAATCCATGGAGCTGAGTGACATTTCCAAAT 1184
QY 1121 GAAAACTGCCAGATTAAACAGATATGGCCACTTTTCAAGCCACCACATCAAAATGTAGAGGA 1180
Db 1185 GAAAACTGCCGAATAAACAGATATGGCTACTTTCAGAGCCACCACATCAAAATGTAGAGGG 1244
QY 1181 ATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGTGCCATGGCCTGAA 1240
Db 1245 ATCTGGAAAGTCAGCATCATGCAGATAGCAGATGTCCCATGCCACACCGCAGCCTGCC 1304
QY 1241 AGTCCCTAATAGACTTTTGTGCTGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACC 1300
Db 1305 AACTCCCTGATGGACTTCACTGTGACCTGCAAGGGGCCACCCCATGGAAGCCTGTACG 1364
QY 1301 ATCATTTCTGACCCACCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTG 1360
Db 1365 ATCATCTCCGACCCACCTGCGAGATGCCAGATGCCAGAACCGGGTCTGCAGCCCTGTGGCTGTG 1424
QY 1361 GATGAGATGTGTCTGTGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTG 1420
Db 1425 GATGGCTGTGCCCTGTCTGTGAGAAAGAGCCTTCAATGGGTCTGGCACCTACTGTGTG 1484
QY 1421 AACCTCACCCCTGGGGATGACACAAGCCTGGCTCTCACGAGCACCTGATTTCTGTTCCT 1480
Db 1485 AATTTCACCTCTGGAGATGATGCAAGCCTGGCCCTCACGAGCACCTGATCTCTATCCCT 1544
QY 1481 GACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCCCTGATCTCCGTGGCTGC 1540
Db 1545 GGCAAGACCCAGACTCCCTCTGAGAGCAGTGAATGGTCTCTGATCTCCATCGGCTGC 1604
QY 1541 TTGGCCATANTTGTCACTGTGATCTCCCTCTTGGTGTAACAAAAACACAGGAATACAAC 1600
Db 1605 CTGGCTGTGCTGTACCATGGTTACCATTTGCTGTACAAAAACACAGGCGTACAAG 1664
QY 1601 CCAATAGAAATAGTCTCTGGGAATGTGTCAGAAAGCAAGGCCTGAGTGTCTTCTCAAC 1660
Db 1665 CCAATAGGAAACTGCCCGAGAACACCGTCAAGGCAAGGCAAGGCGTACAAG 1664
QY 1661 CGTGCAAAAGCCGTGTTCTTCCCGGGAACACAGGAAAGGATCCGCTACTCAAAACCAA 1720
Db 1725 CACGCGAAAGCCCGTCTTCCGAGGAGACGAGGAGAAAGGATCCATTTGCTCCAGGACAAG 1784
QY 1721 GAATTTAAAGGAGTTTCTTAAATTTTCGACCTTGTCTTGAAGCTCACTTTTCAAGTCCCAT 1780
Db 1785 CCAAGGACACTCTAAGTCTTTGGCCCTTCCCTCTGACGAGAACCCACTCTCTGTG-CAT 1843
QY 1781 TGATGTGAGATGTGCTGGA---GTGGCTATTAACCTTTTTCCTAAAGATATTGTTA 1836
Db 1844 GTATGTGAGCTGTGCAGAAAGTATGTGGCTGGGAACCTGTTGTTCTCTAAGGATATTGTAA 1903
QY 1837 AATAGATATTGTGTTTGGGGAAGTTGAATTTTATAGGTTAAA 1881
Db 1904 AATGTATATCGTGGCTTAGGGAGTGTGGTTAAATAGCATTTTGA 1948

RESULT 4

BX397314/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BX397314 1201 bp mRNA linear EST 13-MAY-2003
BX397314 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CS0DI033YF05 3-PRIME, mRNA sequence.
BX397314
BX397314
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4396.f For
more information about this cluster, see

cgi-bin/cluster.cgi?seq=CS0DI033CC03NP1&cluster=4396.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
<http://fulllength.invitrogen.com/> InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DI033CC03NP1.

FEATURES

source

```

1. .1201
/organism="Homo sapiens"
/mol_type="mrna"
/db_xref="taxon:9606"
/clone="CS0DI033YF05"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (
primer. Five prime end enriched, double-strand cDNA
digested with Not I and cloned into the Not I and Eco
sites of the pCMVSPORT 6 vector. Library was normali

```

Query Match	38.5%;	Score 1025;	DB 13;	Length 1201;
Best Local Similarity	95.9%;	Pred. No. 7.9e-156;		
Matches 1072;	Conservative 16;	Mismatches 26;	Indels 4;	Gaps 4;

QY	1499	CCTTTAAGGATGGCAAAACAGTCGCCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACT	1555
Db	1115	CCCTTTARGAKTGCAAAACACGCGCTGATCTCTG-TGGCTGCTTGGCCATATTTGTCACT	1057
QY	1559	GTGATCTCCCTCTTGGTGACAAAAACACAAGGAATACAACCCCAATAGAAAAATAGTCCT	1618
Db	1056	GTGATCTCCCTC-TGGTGACAAAAACACAARGAAWACAACCCCAATAGAAAAATAGTCCT	998
QY	1619	GGGAATGTGGTCAGAAGCAAAAGGCCTGAGTGTCTTTCTC-AACCGTGCAAAAGCGGTGT	1677
Db	997	GGGAATGTGGTCAGAAGCAAAAGGCCTGAGTGTCTGTCTCAACCGTGCAAAAGCGGTGT	938
QY	1678	CTTCCCGGGAACAGGAAAAAGGATCCGCTACTCAAAAAACCAAGAATTTTAAAGGAGTTTC	1737
Db	937	CTTCCCGGGAACAGGAAAAAGGATCCGCTACTCAAAAAACCAAGAATTTTAAAGGAGTTTC	878
QY	1738	TTAAATTCGACCTTGTTTCTGAAGCTCACTTTTCAGTGCCATTGATGTGAGATGTGCTG	1797
Db	877	TTAAATTCGACCTTGTTTCTGAAGCTCACTTTTCAGTGCCATTGATGTGAGATGTGCTG	818
QY	1798	GAGTGGCTATTAAACCTTTTCTTCTTAAAGATTATTGTTAAATAGATATTGTGGTTTGGGG	1857
Db	817	GAGTGGCTATTAAACCTTTTCTTCTTAAAGATTATTGTTAAATAGATATTGTGGTTTGGGG	758
QY	1858	AAGTTGAATTTTATAGGTTAAATGCTCATTTTAGAGATGGGGAGAGGATTATACTGCA	1917
Db	757	AAGTTGAATTTTATAGGTTAAATGCTCATTTTAGAGATGGGGAGAGGATTATACTGCA	698
QY	1918	GGCAGCTTCAGCCATGTTGTGAAACTGATAAAAGCAACTTAGCAAGGCTCTTTTTCATTA	1977
Db	697	GGCAGCTTCAGCCATGTTGTGAAACTGATAAAAGCAACTTAGCAAGGCTCTTTTTCATTA	638
QY	1978	TTTTTTATGTTTCACTTATAAAGCTTTAGGTAACCTAGTAGGATAGAAACACTGTGTCCCG	2037
Db	637	TTTTTTATGTTTCACTTATAAAGCTTTAGGTAACCTAGTAGGATAGAAACACTGTGTCCCG	578
QY	2038	AGAGTAAGGAGAGAAGCTACTATTGATTAGAGCCTAACCCAGGTTAACTGCAAGAAGAGG	2097
Db	577	AGAGTAAGGAGAGAAGCTACTATTGATTAGAGCCTAACCCAGGTTAACTGCAAGAAGAGG	518
QY	2098	CGGGATACTTTTCAGCTTTTCCATGTAACTGTATGCATAAAGCCCAATGTAGTCCAGTTCTA	2157
Db	517	CGGGATACTTTTCAGCTTTTCCATGTAACTGTATGCATAAAGCCCAATGTAGTCCAGTTCTA	458

Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES

source

Location/Qualifiers
1..2431
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9530038P20"
/db_xref="MGI:2399691"
/db_xref="taxon:10090"
/clone="9530038P20"
/sex="male"
/tissue_type="urinary bladder"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
48..1628
/note="unnamed protein product; glycoprotein (transmembrane) nmb (MGD|MGI:1934765, GB|NM_053110, evidence: BLASTN, 99%, match=2272)
putative"
/codon_start=1
/protein_id="BAC37578.1"
/db_xref="GI:26347859"
translation="MESLCGVLGFLLLAAGLPLOAAKRFRDVLGHEQYPDHMRHNQL
RGWSSDENWEHLYPVWRRGDGRWKDSWEGGRVQAVLTS DSPALVGSNITFVNVLV
PRCKEDANGNIVYEKNCRNDLGLTSLHVNWTAGDDGDWEDGTSRSQHLRFPDRR
PFPRPHGWKWSFVYVHTLGOYFQKLGRCSARVSINTVNLTAGPQWMEVTVFRRYGR
AYIPISKVDVYVITDQIPVFTMSQKNDRLNLSDEIFLRDLPIVDVLIDPSPHFLND
SAISYKNPFGDNTGLFVSNHNTLHNTYVINGTFNLNTVQTAVPGPCPPSPSPSTPPP
STPPSPSPPLPTLSTPSPSLMPTGYKSMELSDISNENCRINRYGFRAITIVIGIL
EVSIMQIADVPMPPTQPANSLMDFVTCKGATPMEACTIISDPTCOIAQNRVCSPVAV
DGLCLLSVRRAFNGSGTYCVNFILGDDASLALTSLISIPGKDPDPSPLRAVNGVLISI
GCLAVLVTMTVILLNLG"

ORIGIN

Query Match		35.5%;	Score 943.8;	DB 11;	Length 2431;
Best local similarity		75.9%;	Pred. No. 6.9e-143;		
Matches 1215;		Conservative	0;	Mismatches 337;	Indels 49; Gaps 2;
QY	30	CCTTGGTGCCTCGTCCGTGAGAAATTCAGATGGAATGTCTCTACTATTTCCTGGGATTT	89		
Db	18	CCTGACTGGTTGCAGGCGCTCGGAGTCAGCATGGAAGTCTCTCGCGGGTCTCGGATTT	77		
QY	90	CTGCTCCTGGTGCAGGATTCGCCACTTGATGCGGCCAAACGATTTTCATGATGCTGGGC	149		
Db	78	CTGCTGCTGGTGCAGGACTGCCCTCTCCAGGCTGCCAAGCATTTTCGTGATGCTGGGC	137		
QY	150	AATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGGTCTTCTGAT	209		
Db	138	CATGAACAGTATCCCGATCACAATGAGAGAGCACAAACCAATTACGTGGTGGTCTTCGGAT	197		
QY	210	GAAAATGACTGGAATGAAAACCTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGGAAA	269		
Db	198	GAAAATGAATGGGATGAACACCTGTATCCAGTGTGGAGGAGGGAGCGCAGGTGGAAG	257		
QY	270	AACTCCTGGAAGGAGGCGGTGTGACGGCGTCTGACCAAGTCACTCACCAGCCCTCGTG	329		
Db	258	GACTCCTGGGAAGGAGGCGGTGTGACGGCAGTCTTGACCAAGTCACTCACCAGTCTGGTG	317		
QY	330	GGCTCAATATACATTTGCGGTGAACCTGATATCCCTAGATGCCAAAAGGAAGATGCC	389		
Db	318	GGTTCCAATATCACTTTTGTGGTGAACCTGGTGTTCGCCAGATGCCAAGGAAGATGCT	377		

QY	390	AATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATCCA	449		
Db	378	AATGGCAATATCGTCTATGAGAAGAACTGCAGGAATGATTTGGACTGACATCTGACCTG	437		
QY	450	TATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGACGGGGAATAATGGCACCGGCCAA	509		
Db	438	CATGCTACAACCTGGACTGCGGGGCAGATGATGGTACTGGGAAGATGGCACCGCCGA	497		
QY	510	AGCCATCATAACTGTTCCCTGATGGGAAACCTTTTCTCACACCCCGGATGGAGAGA	569		
Db	498	AGCCAGCATCTCAGGTTCCCGGACAGGAGGCCCTTCCCTCGCCCCCATGGATGGAAGAAA	557		
QY	570	TGGAATTTCACTACGTTCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGATGT	629		
Db	558	TGGAGCTTTGTCTACGTTCTTACACACTTGGCCAGTATTTCCAAAAACTGGGTCGGTGT	617		
QY	630	TCAGTGAGAGTTTCTGTGAACACACAGCCCAATGTGACACTTTGGGCCCTCAACTCATGGAAGTG	689		
Db	618	TCAGCACGGGTTTCTATAAACACAGTCAACTTTCAGACAGTGGCCCTCAGGTCATGGAAGTG	677		
QY	690	ACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGTAC	749		
Db	678	ACTGTCTTTCGAAGATACGGCCGGGCATACATTCATCTCGAAGGTGAAGATGTGTAT	737		
QY	750	GTGGTAACAGATCAGATTCCTGTGTTTGTGACTATGTTCCAGAAAGATCGAATTCATCA	809		
Db	738	GTGATAACAGATCAGATCCCTGTATTTCGTGACCATGTCCAGAAAGATGACAGGAACCTTG	797		
QY	810	TCCGACGAAACCTTCC - CAAAGATCTCCCAATTATGTTTGTATGTCTGATTCATGATCCT	868		
Db	798	TCTGATGAGATCTTCCCTCAGAGACCTCCCCATCGTCTTCGATGTCTTCATTCATGATCCC	857		
QY	869	AGCCACTTCTCAATTATTCTACCATTAACATAAGTGGAGCTTCGGGGGATAATACTGGC	928		
Db	858	AGCCACTTCTCTCAACGACTCTGCCATTTCTTCAAGTGGAACCTTTGGGGACAACACTGGC	917		
QY	929	CTGTTTGTTCACCAATCATACTGTGAATCACACGTATGTCTCAATGGAACCTTCAGC	988		
Db	918	CTGTTTGTCTCCAAATCACACCTTTGAATCACACTTATGTCTCAATGGAACCTTCACAC	977		
QY	989	CTTAACCTCACTGTGAAAGCTGCAGCACACGACCTTG-----	1026		
Db	978	CTTAACCTCACCGTGCAAACTGCAGTGCCCGGGCCATGCCCTCCCCCTTCGCTTCGACT	1037		
QY	1027	-----TCCGCCACCGCCACCACCCACCCAGACCTTCAAAA	1060		
Db	1038	CCGCCCTCCACCTTCAACTCCGCCCTCACCTCCGCCCTCACCTCTGCCACATATCAACA	1097		
QY	1061	CCCACCCCTTCTTTAGGACCTGTGTGTGACAAACCCCTGGAGCTGATGAGGATTCCTGAT	1120		
Db	1098	CCTAGCCCTCTTTAATGCCTACTGTGTACAAATCCATGGAGCTGAGTGACATTTCCAAT	1157		
QY	1121	GAAAACCTGCCAGATTAAACAGATATGSCCACTTTCAAGCCACCATCACAATTGTAGAGGA	1180		
Db	1158	GAAAACCTGCCGAATAAACAGATATGGCTACTTCAGAGCCACCATCACAATTGTAGAGGG	1217		
QY	1181	ATCTTAGAGGTTAACATCATCCAGATGACAGACGTCTGTATGCGGTGCCATGGCCTGAA	1240		
Db	1218	ATCCTGGAAGTCAGCATCATGAGATAGCAGATGTCCCATGCCACACCGCAGCCTGCC	1277		
QY	1241	AGTCCCTAATAGACTTTTGTCTGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACC	1300		
Db	1278	AACCTCCTGATGGACTTCACTGTGACCTGCAAGGGGCCACCCCATGGAAGCCTGTACG	1337		
QY	1301	ATCATTTCTGACCCCACTGCGAGATCACCCAGAACACAGTCTGACGCCCTGTGGATGTG	1360		
Db	1338	ATCATCTCCGACCCCACTGCGAGATGCCAGAACCCGGTCTGACGCCCTGTGGCTGTG	1397		
QY	1361	GATGAGATGTGTCTGTGACTGTGAGACGAACCTTCAATGGTCTTGGGACGTACTGTGTG	1420		
Db	1398	GATGGCTGTGCCTGTGTCTGTGAGAAAGACCTTCAATGGGTCTGGCACCTACTGTGTG	1457		

Y	1421	AACCTCACCTGGGGATGACAAAGCCCTGGCTCTCACGAGACCCCTGATTTCTGTTCCT	1480
b	1458	AATTTCACTCTGGGAGATGATGCAAGCCTGGCCCTCACGAGACCCCTGATCTCTATCCCT	1517
Y	1481	GACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCTGATCTCCGTTGGCTGC	1540
b	1518	GGCAAGACCCAGACTCCCTCTGAGAGCAGTGAATGGTGTCTGATCTCCATCGGCTGC	1577
Y	1541	TTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTAACA	1581
b	1578	CTGGCTGTGCTTGTCAACCATGGTTACCATCTTGTGCTGTACA	1618
RESULT 7			
BX423077/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
BX423077			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			

Matches	949;	Conservative	6;	Mismatches	18;	Indels	3;	Gaps	2;
QY	897	ACTACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTGTTTCCACCAATCATACTGTGA	956						
Db	57	AATTCCCGGATGCTTCGGGGATAATACTGGCCTGTTGTTTCCACCAATCATACTGTGA	116						
QY	957	ATCACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCAC	1016						
Db	117	ATCACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCAC	176						
QY	1017	CAGGACCTTGTCCGCCACCGCCACCAACCCAGACCTTCAAAAACCCACCCCTTCTTTAG	1076						
Db	177	CAGGACCTTGTCCGCCACCGCCACCAACCCAGACCTTCAAAAACCCACCCCTTCTTTAG	236						
QY	1077	GACCTGCTGGTGACAAACCCCTCGAGCTGAGTAGGATTCCTGATGAAAACTGCCAGATTA	1136						
Db	237	GACCTGCTGGTGACAAACCCCTCGAGCTGAGTAGGATTCCTGATGAAAACTGCCAGATTA	296						
QY	1137	ACAGATATGGCCACTTTCAAGCCACCATCACAAATTGTAGAGGAATCTTAGAGGTTAACA	1196						
Db	297	ACAGATATGGCCACTTTCAAGCCACCATCACAAATTGTAGAGGAATCTTAGAGGTTAACA	356						
QY	1197	TCATCCAGATGACAGACGTCTCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAATAGACT	1256						
Db	357	TCATCCAGATGACAGACGTCTCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAATAGACT	416						
QY	1257	TTGTCTGTGACCTGCCAAGGGAGCATTTCCCACGGAGGTCTGTACCATCATTTCTGACCCCCA	1316						
Db	417	TTGTCTGTGACCTGCCAAGGGAGCATTTCCCACGGAGGTCTGTACCATCATTTCTGACCCCCA	476						
QY	1317	CCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGC	1376						
Db	477	CCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGC	536						
QY	1377	TGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGG	1436						
Db	537	TGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGG	596						
QY	1437	ATGACACAAGCCTGGCTCTCACGAGCACCTTGATTTCTGTTCTCGACAGAGACCCAGCCT	1496						
Db	597	ATGACACAAGCCTGGCTCTCACGAGCACCTTGATTTCTGTTCTCGACAGAGACCCAGCCT	656						
QY	1497	CGCCTTTAAGGATGGCAACACAGTCCCTTGATCTCCGTTGGCTGCTTGGCCATATTTGTCA	1556						
Db	657	CGCCTTTAAGGATGGCAACACAGTCCCTTGATCTCCGTTGGCTGCTTGGCCATATTTGTCA	716						
QY	1557	CTGTGATCTCCCTCTTGGTGTAACAAAAACACAAGGAATACAACCCCAATAGAAAATAGTC	1616						
Db	717	CTGTGATCTCCCTCTTGGTGTAACAAAAACACAAGGAATACAACCCCAATAGAAAATAGTC	776						
QY	1617	CTGGGAATGTGGTCAGAAAGCAAGGCCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGT	1676						
Db	777	CTGGGAATGTGGTCAGAAAGCAAGGCCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGT	836						
QY	1677	TCTTCCCGGAAACCCAGGAAAGGATCCGCTACTCAAAAAACCAAGAAATTTAAAGGAGTTT	1736						
Db	837	TCTTCCCGGAAAC - CAGGAAAGGATCCGCTACTCAAAAAACCAAGAAATTTAAAGGAGTTT	895						
QY	1737	CTTAAATTTCCAGCTTGTTTTCTGAAGCTCACTTTTCAGTGCCTTGTGATGTGAGATGTGCT	1796						
Db	896	CTTAAATTTCCAGCTTGTTTTCTGAAGCTCACTTTTCAGTKCCATTGTGTKAGAGGTGCK	955						
QY	1797	GGAGTGGCTATTAAACCTTTTTTTCCTAAAGATTATTGTTAAATAGATATTGTGTTTGGG	1856						
Db	956	GGAGTGGCTATTAAACCTTTTTTTCCTAAAGATTATTGTTAAATAGATATTGTGTTTGGG	1013						
QY	1857	GAAGTTGAATTTTTTA	1872						
Db	1014	GAGTTGRATTTTWTWTA	1029						

RESULT 10
BX379979

LOCUS	BX379979	1054 bp	mRNA	linear	EST 08-MAY-2003
DEFINITION	BX379979 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI042YI21 5-PRIME, mRNA sequence.				
ACCESSION	BX379979				
VERSION	BX379979.1	GI:30456816			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 1054)				
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4396.f For more information about this cluster, see http://www.genoscope.cns.fr/ cgi-bin/cluster.cgi?seq=CS0DI042AE11QPl&cluster=4396.f. Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS0DI042AE11QPl.				
FEATURES	Location/Qualifiers				
source	1. .1054 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0DI042YI21" /tissue_type="PLACENTA COT 25-NORMALIZED" /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED" /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."				
ORIGIN					
Query Match	34.3%;	Score 914;	DB 13;	Length 1054;	
Best Local Similarity	96.0%;	Pred. No. 6.8e-138;			
Matches	960;	Conservative 2;	Mismatches 2;	Indels 36;	Gaps 1;
QY	897	ACTACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTGTTTCCACCAATCATACTGTGA	956		
Db	55	ATTACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTGTTTCCACCAATCATACTGTGA	114		
QY	957	ATCACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCAC	1016		
Db	115	ATCACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCAC	174		
QY	1017	CAGGACCTTGTCCGCCACCGCCACCAACCCAGACCTTCAAAAACCCACCCCTTCTTT--	1074		
Db	175	CAGGACCTTSTCCGCCACCGCCACCAACCCAGACCTTCAAAAACCCACCCCTTCTTTAG	234		
QY	1075	-----AGGACCTGCTGGTGACAACCCCTGG 1100			
Db	235	CAACTACTCTAAAAATCTTATGATTCAAAACACCCAGGACCTGCTGGTGACAACCCCTGG	294		
QY	1101	AGCTGAGTAGGATTCCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCA	1160		
Db	295	AGCTGAGTAGGATTCCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCA	354		
QY	1161	CCATCACAAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCCTGA	1220		
Db	355	CCATCACAAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCCTGA	414		
QY	1221	TGCCGGTGCATGGCCTGAAAGCTCCCTAAATAGACTTTGTGCTGACCTGCCAAGGGAGCA	1280		
Db	415	TGCCGGTGCATGGCCTGAAAGCTCCCTAAATAGACTTTGTGCTGACCTGCCAAGGGAGCA	474		
QY	1281	TTCCCACGGAGGTCTGTACCATCATTTCTGACCCCACTCGGAGATCACCCAGAACACAG	1340		

|||||
Db 475 TTCCCACGGAGTCTGTACCATCATTTCTGACCCACCTGCGAGATACCCAGAACACAG 534
QY 1341 TCTGCAGCCCTGTGGATGGATGAGATGTGTCTGTCTGCTGACTGTGAGACGAACCTTCAATG 1400
Db 535 TCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTCTGCTGACTGTGAGACGAACCTTCAATG 594
QY 1401 GGTCTGGGACGTACTGTGTGAACCTCAACCTGGGGGATGACACAAGCCTGGCTCTCACGA 1460
Db 595 GGTCTGGGACGTACTGTGTGAACCTCAACCTGGGGGATGACACAAGCCTGGCTCTCACGA 654
QY 1461 GCACCCCTGATTTCTGTTCCCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACACAGTG 1520
Db 655 GCACCCCTGATTTCTGTTCCCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACACAGCG 714
QY 1521 CCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTA 1580
Db 715 CCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTA 774
QY 1581 AAAACACAAAGGAATACAAACCCCAATAGAAAATAGTCTCGGAATGTGGTCAGAAGCAAG 1640
Db 775 AAAACACAAAGGAATACAAACCCCAATAGAAAATAGTCTCGGAATGTGGTCAGAAGCAAG 834
QY 1641 GCCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTTCTTCCCGGAAACCCAGGAAAAGG 1700
Db 835 GCCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTTCTTCCCGGAAACCCAGGAAAAGG 894
QY 1701 ATCCGCTACTCAAAAACCAAGAAATTAAGGAGTTTCTTAAATTTTCGACCTTGTCTGA 1760
Db 895 ATCCGCTACTCAAAAACCAAGAAATTAAGGAGTTTCTTAAATTTTCGACCTTGTCTGA 954
QY 1761 AGCTCACTTTTCAGTGCCATGATGTGAGATGTCTGGAGTGGCTATTAAACCTTTTTC 1820
Db 955 AGCTCACTTTTCAGTGCCATGATGTGAGATGTCTGGAGTGGCTATTAAACCTTTTTC 1014
QY 1821 CTAAAGATTATTGTTAAATAGATATTGTGGTTTGGGGAAG 1860
Db 1015 CTAAAGATTATTGTTAAATAGATATTGTGGTTTGGGGRAG 1054

RESULT 11
BX336884
LOCUS
DEFINITION
BX336884 Homo sapiens PLACENTA COT 25-NORMALIZED EST 01-MAY-2003
clone CS0DI033YF05 5-PRIME, mRNA sequence.
BX336884
VERSION
BX336884.1 GI:30308517
KEYWORDS
SOURCE
EST.
Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1201)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4396.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DI033CC03QP1
&cluster=4396.f. Contact : Feng Liang Email : fliang@lifetech.com
URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DI033CC03QP1.

FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI033YF05"

/tissue type="PLACENTA COT 25-NORMALIZED"
/clone lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and Ecor V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 34.2%; Score 910.6; DB 13; Length 1201;
Best Local Similarity 99.2%; Pred. No. 2.2e-137;
Matches 935; Conservative 1; Mismatches 5; Indels 2; Gaps 2;
QY 28 AACCTTGGTGCCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCCCTGGGAT 87
Db 103 ACCTTGAGTGCCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCCCTGGGAT 162
QY 88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGTATGCCGCCAAACGATTTTCATGATGTCTGG 147
Db 163 TTCTGCTCCTGGCTGCAAGATTGCCACTTGTATGCCGCCAAACGATTTTCATGATGTCTGG 222
QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGGTCTCTG 207
Db 223 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGGTCTCTG 282
QY 208 ATGAAATGACTTGAATGAAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 267
Db 283 ATGAAATGACTTGAATGAAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 342
QY 268 AAAACTCCTGGAAGGGAGGCCGTGTGCAGGCGGTCTGACCCAGTGTGACCCAGTGTGAC 402
Db 343 AAAACTCCTGGAAGGGAGGCCGTGTGCAGGCGGTCTGACCCAGTGTGACCCAGTGTGAC 327
QY 328 TGGGCTCAAATATAACATTTGCGGTGAACCTGTATTCCTAGATGCCAAAAGGAAGATG 387
Db 403 TGGGCTCAAATATAACATTTGCGGTGAACCTGTATTCCTAGATGCCAAAAGGAAGATG 462
QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 447
Db 463 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 522
QY 448 CATATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGCAGGGGAAAATGGCACCGGCC 507
Db 523 CGTATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGCAGGGGAAAATGGCACCGGCC 582
QY 508 AAAGCCATCATAAACGTCTTCCCTGTATGGGAAACCTTTTCTCACCCCGGATGGAGAA 567
Db 583 AAAGCCATCATAAACGTCTTCCCTGTATGGGAAACCTTTTCTCACCCCGGATGGAGAA 642
QY 568 GATGGAATTTTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGAT 627
Db 643 GATGGAATTTTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGAT 702
QY 628 GTTCAGTGAGAGTTTCTGTGAACACAGCCAATGTGACACTTGGGCGCTCAACTCATGGAAG 687
Db 703 GTTCAGTGAGAGTTTCTGTGAACACAGCCAATGTGACACTTGGGCGCTCAACTCATGGAAG 762
QY 688 TGACTGTCTACAGAAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAAGATGTGT 747
Db 763 TGACTGTCTACAGAAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAAGATGTGT 822
QY 748 ACGTGTAAACAGATCAGATTCCCTGTGTTTGTGACTATGTTCCAGAAAGACGATCGAAATT 807
Db 823 ACGTGTAAACAGATCAGATTCCCTGTGTTTGTGACTATGTTCCAGAAAGACGATCGAAATT 882
QY 808 CATCCGACGAAACCTTCC-CAAAGATCTCCCCATTATGTTTGTATGTCCTGATTCATGATC 866
Db 883 CATCCGACGAAACCTTCC-CAAAGATCTCCCCATTATGTTTGTATGTCCTGATTCATGATC 942
QY 867 CTAGCCACTTCCCTCAATTATTCTACCAATTAACATAAGTGGAGCTTCGGGGATAATACTG 926
Db 943 CTAGCCACTTCCCTCAATTATTCTACCAATTAACATAAGTGGAGCTTCGGGGATAATACTG 1002
QY 927 GCCTGTTTGTTCACCAATCATACTGTGAATCACACGATGTGT 969

Db 1003 GCCTGTTGTTT-CMCCAATCATACTGTGAATCACACGTATGT 1044

RESULT 12
BX364871 1201 bp mRNA linear EST 05-MAY-2003
LOCUS
DEFINITION BX364871 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA clone CS0DI024YA06 5-PRIME, mRNA sequence.

ACCESSION BX364871 GI:30380860
VERSION
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 1201)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4396.f For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS1AI006ZG04QP1&cluster=4396.f. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CS1AI006ZG04QP1.

FEATURES
source Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DI024YA06"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 34.1%; Score 906.4; DB 13; Length 1201;
Best Local Similarity 93.8%; Pred. No. 1.1e-136;
Matches 959; Conservative 23; Mismatches 34; Indels 6; Gaps 4;

QY 28 AACCTTGGTGCCTGCGTCGAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGGAT 87
Db 94 ACCTTGAGTGCCTGCGTCGTCGAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGGAT 153
QY 88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTCATGATGTGCTGG 147
Db 154 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTCATGATGTGCTGG 213
QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTAATGGCTGCTCTCTG 207
Db 214 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTAATGGCTGCTCTCTG 273
QY 208 ATGAAAATGACTGGAATGAAAACCTTACCCAGTGTGGAGCGGGAGACATGAGGTGGA 267
Db 274 ATGAAAATGACTGGRATGAAAACCTTACCCAGTGTGGAGCGGGAGACATGAGGTGGA 333
QY 268 AAAACTCCTGGAAGGAGCGCGTGTGCAGCGGCTCCTGACCAGTGACTCACCAGCCCTCG 327
Db 334 AAAACTCCTGGAAGGAGCGCGTGTGCAGCGGCTCCTGACCAGTGACTCACCAGCCCTCG 393
QY 328 TGGGCTCAAATATAACATTTCGGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG 387
Db 394 TGGGCTCAAATATAACATTTCGGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG 453

QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 447
Db 454 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 513
QY 448 CATATGTTTACAACTGGACAGCATGGTTCAGAGGACAGTGCACGGGGAATAATGGCACCGGC 507
Db 514 CGTATGTTTACAACTGGACAGCATGGTTCAGAGGACAGTGCACGGGGAATAATGGCACCGGC 573
QY 508 AAAGCCCATATAACGTCCTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAA 567
Db 574 AAAGCCCATATAACGTCCTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAA 633
QY 568 GATGGAATTTTCATCTACGTCCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGAT 627
Db 634 GATGGAATTTTCATCTACGTCCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGAT 693
QY 628 GTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG 687
Db 694 GTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG 753
QY 688 TGACTGTCTACAGAAAGACATGGACGGGCATATGTTCCCATCGCAAGTGAAGATGTGT 747
Db 754 TGACTGTCTACAGAAAGACATGGACGGGCATATGTTCCCATCGCAAGTGAAGATGTGT 813
QY 748 ACGTGGTAAACAGATCAGATTCTCTGTGTTTGTGACTATGTTCCAGAAACGATCGAAATT 807
Db 814 ACGTGGTAAACAGATCAGATTCTCTGTGTTTGTGACTATGTTCCAGAAACGATCGAAATT 873
QY 808 CATCCGACGAAACCTTCC-CAAAGATCTCCCCATTATGTTTGTATGTCCTGATTCATGATC 866
Db 874 CATCCGACGAAACCTTCC-CAAAGATCTCCCCATTATGTTTGTATGTCCTGATTCATGATC 933
QY 867 CTAGCCACTTCTCTCAATATTCTACCAATTAACATAAGTGGAGCTTCGGGGATAATACTG 926
Db 934 CTAGCCACTTCTCTCAATATTCTACCAATTAACATAAGTGGAGCTTCGGGGATAATACTG 993
QY 927 GCCTGTTTGTTCACCAATCATACTGTGAATCACACGATATGTGCTCAATGGAACTTCA 986
Db 994 KGCYGTGTTTGTTCCTCA-CAATCACTGTGAATCACACACKTATTTCTCMATKG--AACYTCA 1050
QY 987 GCCTTAACCTCACTGTGAAAGCTGCAGCACCAAGACCTTGTCCGCCACCCGCCACCAAC 1046
Db 1051 GCCTTACCYYCMGTGK--AAGSTGCMGCCAGGACTTTTCTSCMCCGWCWCMCCCARA 1108
QY 1047 CC 1048
Db 1109 CY 1110

RESULT 13
BX458449 1012 bp mRNA linear EST 22-MAY-2003
LOCUS
DEFINITION BX458449 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE003YI18 5-PRIME, mRNA sequence.
ACCESSION BX458449
VERSION BX458449.1 GI:31025019
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1012)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 4396.f For more information about this cluster, see

Db 541 ||||| GACCCACCTGCGAGATCACCCAGAAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATG 600

QY 1370 TGTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACC 1429

Db 601 TGTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACC 660

QY 1430 CTGGGGGATGACACAAGCCTGGGTCTCACGAGCACCCCTGATTTCTGTTCCCTGACAGAGAC 1489

Db 661 CTGGGGGATGACACAAGCCTGGGTCTCACGAGCACCCCTGATTTCTGTTCCCTGACAGAGAC 720

QY 1490 CCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCCTGATCTCCGTTGGCTGCTTGGCCATA 1549

Db 721 CCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCCTGATCTCCGTTGGCTGCTTGGCCATA 780

QY 1550 TTTGTCACTGTGATCTCCCTCTTGGTGTAACAAAAACACAAGGAATACAACCCCAATAGAA 1609

Db 781 TTTGTCACTGTGATCTCCCTCTTGGTGTAACAAAAACACAAGGAATACAACCCCAATAGAA 840

QY 1610 AATAGTCCTGGGAATGTGCTCAGAAGCAAAGG-CCTGAGTGTCTTTCTCAACC 1661

Db 841 AATAGTCCTGGGAATGTGCTCAGAACAAGGCGCCTGAGTGTCTTTCTCAACC 893

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 10, 2004, 13:27:49 ; Search time 218 Seconds
(without alignments)
7702.697 Million cell updates/sec

Title: US-10-039-272-1
Perfect score: 4850
Sequence: 1 cggcagcagggccagagga.....aaaaaaaaaaaaaaaaaaaaa 2661

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTC_spool_p/US10039272/runat_07092004_144223_21629/app_query.fasta_1.2823
-DB=SPTREMBL_25 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10039272 @CGN_1_1_296 @runat_07092004_144223_21629 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	ID Description
1	2983	61.5	572 4 Q8N1A1 Q8n1a1 homo sapien

2	2170	44.7	574	11	Q99P91	Q99p91 mus musculu
3	2170	44.7	574	11	Q9QXA0	Q9qxa0 mus musculu
4	2170	44.7	574	11	Q8BVV9	Q8bvv9 mus musculu
5	2166	44.7	574	11	Q8BXL4	Q8bxl4 mus musculu
6	2136	44.0	572	11	Q9QZF6	Q9qzf6 rattus norv
7	2015	41.5	526	11	Q8BVA0	Q8bva0 mus musculu
8	1029	21.2	206	4	Q96F58	Q96f58 homo sapien
9	1022	21.1	206	4	Q8IXJ5	Q8ixj5 homo sapien
10	644.5	13.3	626	11	Q9CZB2	Q9czb2 mus musculu
11	372	7.7	461	6	Q97884	Q97884 equus cabal
12	271.5	5.6	270	13	Q93391	Q93391 coturnix co
13	247.5	5.1	236	11	Q9QY67	Q9qy67 mus musculu
14	198.5	4.1	423	4	Q8N0W9	Q8n0w9 homo sapien
15	198.5	4.1	435	4	Q8N3G9	Q8n3g9 homo sapien
16	192	4.0	397	4	Q8N3R2	Q8n3r2 homo sapien
17	167	3.4	141	11	Q9QY70	Q9qy70 mus musculu
18	151	3.1	354	4	Q8IY46	Q8iy46 homo sapien
19	141.5	2.9	906	17	Q8TPY9	Q8tpy9 methanosarc
20	135	2.8	446	10	O22458	O22458 chlamydomon
21	134.5	2.8	1817	17	Q8TI59	Q8ti59 methanosarc
22	131	2.7	879	17	Q8PWJ6	Q8pwj6 methanosarc
23	128.5	2.6	688	17	Q8TR88	Q8tr88 methanosarc
24	127.5	2.6	3988	17	Q8TPZ1	Q8tpz1 methanosarc
25	126	2.6	603	16	Q7U5X8	Q7u5x8 synechococc
26	121.5	2.5	525	11	Q99PS8	Q99ps8 rattus norv
27	121.5	2.5	1188	17	Q8TPZ7	Q8tpz7 methanosarc
28	121	2.5	1081	6	Q9BE73	Q9be73 macaca fasc
29	120	2.5	766	4	P82987	P82987 homo sapien
30	119.5	2.5	507	16	Q8FY77	Q8fy77 brucella su
31	119.5	2.5	510	11	Q9ESB2	Q9esb2 rattus norv
32	119.5	2.5	2016	17	Q8TUS9	Q8tjs9 methanosarc
33	119	2.5	1948	17	Q8TI51	Q8ti51 methanosarc
34	118.5	2.4	2029	17	Q8TI44	Q8ti44 methanosarc
35	118	2.4	2869	6	Q28398	Q28398 erinaceus e
36	117	2.4	883	17	Q8TTT9	Q8ttt9 methanosarc
37	116.5	2.4	703	11	Q8BGJ1	Q8bgj1 mus musculu
38	116	2.4	1698	17	Q8TPK7	Q8tpk7 methanosarc
39	114.5	2.4	515	11	Q99PS7	Q99ps7 rattus norv
40	114.5	2.4	594	5	Q9VEP4	Q9vep4 drosophila
41	114	2.4	2395	5	Q27167	Q27167 paramecium
42	113.5	2.3	525	11	Q99PS6	Q99ps6 mus musculu
43	113.5	2.3	525	11	Q9ESB3	Q9esb3 mus musculu
44	113.5	2.3	525	11	Q99PS5	Q99ps5 mus musculu
45	113	2.3	475	10	Q949Z1	Q949z1 arabidopsis

ALIGNMENTS

RESULT 1
Q8N1A1
ID Q8N1A1 PRELIMINARY; PRT; 572 AA.
AC Q8N1A1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to glycoprotein (Transmembrane) nmb.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Lung;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC032783; AAH32783.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR000601; PKD.
DR Pfam; PF00801; PKD; 1.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS50093; PKD; 1.
KW Transmembrane.
SQ SEQUENCE 572 AA; 63922 MW; 2465C12C2F0F3996 CRC64;

Alignment Scores:				
Pred. No.:	1.45e-281	Length:	572	
Score:	2983.00	Matches:	559	
Percent Similarity:	97.73%	Conservative:	0	
Best Local Similarity:	97.73%	Mismatches:	1	
Query Match:	61.51%	Indels:	13	
DB:	4	Gaps:	1	
US-10-039-272-1 (1-2661) x Q8N1A1 (1-572)				
QY	60	ATGGAATGTTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT	119	
Db	1	MetGluCysLeuTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp	20	
QY	120	GCCGCCAAACGATTTCATGATGTGCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAG	179	
Db	21	AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40	
QY	180	CACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTCTACCCA	239	
Db	41	HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60	
QY	240	GTGTGGAAGCGGGAGACATGAGGTGGAATACTCTGGAAGGAGGCCGTGTGCAGCG	299	
Db	61	ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla	80	
QY	300	GTCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAAATATAACATTTGCGGTGAACCTG	359	
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100	
QY	360	ATATTCCTTAGATGCCAAAAGGAAGATGCCAATGCAACATAGTCTATGAGAAGAACTGC	419	
Db	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120	
QY	420	AGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGGTCAGAG	479	
Db	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140	
QY	480	GACAGTGACGGGAAATGGCACCGGCCAAAGCCATCATAACTCTTCCCTGATGGGAAA	539	
Db	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys	160	
QY	540	CCTTTTCCTCACCCCGGATGGAGAAGATGGAATTCATCTACGTCTTCCACACACTT	599	
Db	161	ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu	180	
QY	600	GGTCAGTATTTCCAGAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT	659	
Db	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200	
QY	660	GTGACACTTGGGCCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATAT	719	
Db	201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220	
QY	720	GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTGTG	779	
Db	221	ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal	240	
QY	780	ACTATGTTCCAGAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC	838	
Db	241	ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro	260	
QY	839	ATTATGTTTGATGTCCTGATTCATGATCCTAGCCACTTCCCTCAATATTCTACCATTAAC	898	
Db	261	IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280	
QY	899	TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATACTGTGAAT	958	
Db	281	TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300	
QY	959	CACACGTATGTGTTCAATGGAACCTTCAGCCTTAACCTCAGTGTGAAAGCTGCAGCACCA	1018	

Db	301	HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro	320
QY	1019	GGACCTTGTCGCCACCGCCACACCCAGACCTTCAAAACCCACCCCTTCTTTA---	1075
Db	321	GlyProCysProProProProProProProProProProProProProProProProPro	340
QY	1076	-----GGACCTGCTGCTGACAAACCCCTGGAG	1102
Db	341	ThrThrLeuLysSerTyrAspSerAsnThrProGlyProAlaGlyAspAsnProLeuGlu	360
QY	1103	CTGAGTAGGATTCTCTGATGAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACC	1162
Db	361	LeuSerArgIleProAspGluAsnCysGlnIleAsnArgTyrGlyHisPheGlnAlaThr	380
QY	1163	ATCACAATTGTAGAGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCTCTGATG	1222
Db	381	IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet	400
QY	1223	CCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTTGTCTGACCTGCCAAGGGAGCATT	1282
Db	401	ProValProTrpProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle	420
QY	1283	CCCACGGAGGTCTGTACCATCATTTCTGACCCCCACCTGGGAGATCACCCAGAACACAGTC	1342
Db	421	ProThrGluValCysThrIleIleSerAspProThrCysGluIleThrGlnAsnThrVal	440
QY	1343	TGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTCTGACTGTGAGACGAACTTCAATGGG	1402
Db	441	CysSerProValAspValAspGluMetCysLeuLeuThrValArgArgThrPheAsnGly	460
QY	1403	TCTGGGACGTACTGTGTGAACCTCACCCCTGGGGATGACACAAGCCTGGCTCTCAGCAGC	1462
Db	461	SerGlyThrTyrCysValAsnLeuThrLeuGlyAspAspThrSerLeuAlaLeuThrSer	480
QY	1463	ACCCTGATTTCTGTTCTCTGACAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCC	1522
Db	481	ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla	500
QY	1523	CTGATCTCCGTTGGCTGTTGGCCATATTTGTCTACTGTGATCTCCCTCTTGGTGACAAA	1582
Db	501	LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuLeuValTyrLys	520
QY	1583	AAACACAAGGAATACACCCAATAGAAAATAGTCTCTGGGAATGTGGTCAGAACCAAGGC	1642
Db	521	LysHisLysGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerLysGly	540
QY	1643	CTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCGGGAAACCCAGGAAACGAT	1702
Db	541	LeuSerValPheLeuAsnArgAlaLysAlaValPhePheProGlyAsnGlnGluLysAsp	560
QY	1703	CCGCTACTCAAAACCAAGAAATTTAAAGGAGTTTCT	1738
Db	561	ProLeuLeuLysAsnGlnGluPheLysGlyValSer	572
RESULT 2			
Q99P91	ID	Q99P91	PRELIMINARY; PRT; 574 AA.
AC	Q99P91;		
DT	01-JUN-2001	(TrEMBLrel. 17, Created)	
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)	
DT	01-OCT-2003	(TrEMBLrel. 25, last annotation update)	
DE	Dendritic cell-associated transmembrane protein (Glycoprotein (Transmembrane) nmb).		
GN	GPNMB OR DCHIL.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=BALB/c;		
RA	Shikano S., Bonkobara M., Zukas P.K., Ariizumi K.;		
RT	"Molecular Cloning of a Dendritic Cell-Associated Transmembrane		

RT Protein, DC-HIL, that Promotes RGD-Dependent Adhesion of Endothelial
RL Cells Through Recognition of Heparan Sulfate Proteoglycans.";
RN Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF322054; AAK14240.1; -;
DR EMBL; BC026375; AAH26375.1; -;
DR MGD; MGI:1934765; Gpmb.
DR GO; GO:0005887; C:integral to plasma membrane; IDA.
DR GO; GO:0005194; F:cell adhesion molecule activity; IDA.
DR GO; GO:0008201; F:heparin binding; IDA.
DR GO; GO:0005178; F:integrin binding; IDA.
DR GO; GO:0007155; P:cell adhesion; IDA.
DR InterPro; IPR000583; GATase_2.
DR InterPro; IPR000601; PKD.
DR Pfam; PF00801; PKD; 1.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
DR PROSITE; PS50093; PKD; 1.
KW Transmembrane.
SQ SEQUENCE 574 AA; 63674 MW; B6AE9AC27AE6ACD0 CRC64;

Alignment Scores:
Pred. No.: 3.05e-202 Length: 574
Score: 2170.00 Matches: 402
Percent Similarity: 82.63% Conservative: 69
Best Local Similarity: 70.53% Mismatches: 83
Query Match: 44.74% Indels: 17
DB: 11 Gaps: 2

US-10-039-272-1 (1-2661) x Q99P91 (1-574)

QY 60 ATGGAATGCTCTACTATTTCCTGGGATTTCTGCTCCTGGTGCAGATTGCCACITGAT 119
Db 1 MetGluSerLeuCysGlyValLeuGlyPheLeuLeuAlaGlyLeuProLeuGln 20
QY 120 GCGCCAAACGATTTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGAG 179
Db 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAsnHisMetArgGlu 40
QY 180 CACAATCAATTAATGGCTGCTTCTGATGAAATAGCTGGAATGAAAACTCTACCCCA 239
Db 41 HisAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGluHisLeuTyrPro 60
QY 240 GTGTGGAAGCGGGAGACATGAGTGGGAAAACTCCTGGAAGGAGGCCGTGTGAGCGG 299
Db 61 ValTrpArgArgGlyAspGlyArgTrpLysAspSerTrpGluGlyArgValGlnAla 80
QY 300 GTCCTGACCAGTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGCTGGTTATCTGCTGATCCATATGTTTACAACCTGACAGCATGGTCAGAG 479
Db 121 ArgAsnAspLeuGlyLeuThrSerAspLeuHisValTyrAsnTrpThrAlaGlyAlaAsp 140
QY 480 GACAGTGACGGGAAATAATGGCAACCGGCCAAAGCCATCATAACGCTCTTCCCTGATGGAAA 539
Db 141 AspGlyAspTrpGluAspGlyThrSerArgSerGlnHisLeuArgPheProAspArgArg 160
QY 540 CCTTTTCCTCACCACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCCTCCACACATT 599
Db 161 PropheProArgProHisGlyTrpLysLysTrpSerPheValTyrValPheHisThrLeu 180
QY 600 GGTCAATTTCCAGAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAAT 659
Db

Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerAlaArgValSerIleAsnThrValAsn 200
QY 660 GTGACACATTGGGCTCAACTCATGGAAGTGAAGTGTCTACAGAAGACATGGACGGCATAT 719
Db 201 LeuThrAlaGlyProGlnValMetGluValThrValPheArgArgTyrGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTGTG 779
Db 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetSerGlnLysAsnAspArgAsnLeuSerAspGluIlePheLeuArgAspLeuPro 260
QY 839 ATTATGTTTGATGTCCTGATTTCATGATCCTAGCCACTTCCTCAATTATTCTACCATTAAC 898
Db 261 IleValPheAspValLeuIleHisAspProSerHisPheLeuAsnAspSerAlaIleSer 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTGTTTCCACCAATCATACTGTGAAT 958
Db 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnHisThrLeuAsn 300
QY 959 CACACGTATGTCCTCAATGGAACCTTCAGCCCTTAACCTCAGTGTGAAAGCTGCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheAsnLeuAsnLeuThrValGlnThrAlaValPro 320
QY 1019 GGACCTTGTCCG-----CCACCGCCACCA 1042
Db 321 GlyProCysProProProSerProSerThrProProProProSerThrProProSerPro 340
QY 1043 CCACCCAGACCT-----TCAAAACCCACCCCTTCTTTAGGACCTGCTGGTGAC 1090
Db 341 ProProSerProLeuProThrLeuSerThrProSerProSerLeuMetProThrGlyTyr 360
QY 1091 AACCCCTGGAGCTGAGTAGGATTCTGTATGAAACCTGCAGATTAAACAGATATGCCAC 1150
Db 361 LysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyr 380
QY 1151 TTTCAAGCCACCATCACAATTTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACA 1210
Db 381 PheArgAlaThrIleThrIleValGluGlyIleLeuGluValSerIleMetGlnIleAla 400
QY 1211 GACGTCCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTTGTGCGTACCTGC 1270
Db 401 AspValProMetProThrProGlnProAlaAsnSerLeuMetAspPheThrValThrCys 420
QY 1271 CAAGGGAGCATTTCCACGGAGGTCTGTACCATCATTTCTGACCCCACTGCGAGATCACCC 1330
Db 421 LysGlyAlaThrProMetGluAlaCysThrIleIleSerAspProThrCysGlnIleAla 440
QY 1331 CAGAACACAGTCTGACGCCCTGTGGATGTGGATGGATGATGTGTCTGTGACTGTGAGACGA 1390
Db 441 GlnAsnArgValCysSerProValAlaValAspGlyLeuCysLeuLeuSerValArgArg 460
QY 1391 ACCTTCAATGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGGATGACACAAGCCTG 1450
Db 461 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeu 480
QY 1451 GCTCTCAGGACACCCCTGATTTCCTGTTCTTCGACAGAGACCCAGCCCTCGCCTTTAAGGATG 1510
Db 481 AlaLeuThrSerThrLeuIleSerIleProGlyLysAspProAspSerProLeuArgAla 500
QY 1511 GCAAACAGTCCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTC 1570
Db 501 ValAsnGlyValLeuIleSerIleGlyCysLeuAlaValLeuValThrMetValThrIle 520
QY 1571 TTGGTGTAACAAAAACACAAGGAATACAAACCAATAGAAAATAGTCTGGGAATGTGGTC 1630
Db 521 LeuLeuTyrLysLysHisLysAlaTyrLysProIleGlyAsnCysProArgAsnThrVal 540
QY 1631 AGAAGCAAGGCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGAAAC 1690
Db 541 LysGlyLysGlyLeuSerValLeuLeuSerHisAlaLysAlaProPhePheArgGlyAsp 560


```
QY 1691 CAGGAAAAGGATCCGCTACTCAAAACCAA 1720
Db |||||||
QY 561 GlnGluLysAspProLeuLeuGlnAspLys 570
Db |||||||

RESULT 3
Q9QXA0
ID Q9QXA0 PRELIMINARY; PRT; 574 AA.
AC Q9QXA0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transmembrane glycoprotein.
GN GPNMB OR NMB.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Bachner D., Schroder D., Gross G.;
RT "Isolation and developmental expression analysis of murine nmb.";
RL Submitted (NOV-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AJ251685; CAB65272.1; -.
DR MGD; MGI:1934765; Gpmb.
DR GO; GO:0005887; C:integral to plasma membrane; IDA.
DR GO; GO:0005194; F:cell adhesion molecule activity; IDA.
DR GO; GO:0008201; F:heparin binding; IDA.
DR GO; GO:0005178; F:integrin binding; IDA.
DR GO; GO:0007155; P:cell adhesion; IDA.
DR InterPro; IPR000583; GATase_2.
DR InterPro; IPR000601; PKD.
DR Pfam; PF00801; PKD; 1.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
DR PROSITE; PS50093; PKD; 1.
KW Transmembrane.
SQ SEQUENCE 574 AA; 63681 MW; E0759D7626F0829A CRC64;

Alignment Scores:
Pred. No.: 3.05e-202 Length: 574
Score: 2170.00 Matches: 402
Percent Similarity: 82.63% Conservative: 69
Best Local Similarity: 70.53% Mismatches: 83
Query Match: 44.74% Indels: 17
DB: 11 Gaps: 2

US-10-039-272-1 (1-2661) x Q9QXA0 (1-574)
QY 60 ATGGAATGCTCTACTATTCTCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db |||||||
QY 120 GCCGCCAAACGATTTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db |||||||
QY 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40
QY 180 CACAATCAATTAAATGGCTGGTCTTCTGATGAAAAATGACTGGAATGAAAACTCTACCCA 239
Db |||||||
QY 41 HisAsnGlnLeuArgGlyTyrSerSerAspGluAsnGluTyrAspGluHisLeuTyrPro 60
QY 240 GTGTGGAACGGGAGACATGAGTGGAAAAAACTCCTGGAAGGAGGCCGTGTGCAGGCG 299
Db |||||||
QY 61 ValTyrArgGlyAspGlyArgTyrLysAspSerTyrGluGlyArgValGlnAla 80
QY 300 GTCCTGACCACTGACTCACCAGCCCTCGTGGCTCAATATATAACATTGCGGTGAACCTG 359
Db |||||||
QY 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAAGGAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db |||||||
QY 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
Db |||||||
```

```
QY 420 AGAAATGAGGCTGTTTATCTCTGATCCATATGTTTACAACATGGACGATGGTCAGAG 479
Db |||||||
QY 121 ArgAsnAspLeuGlyLeuThrSerAspLeuHisValTyrAsnTyrThrAlaGlyAlaAsp 140
Db |||||||
QY 480 GACAGTGACGGGAAAAATGGACCGGCCCAAAGCCATCATTAACGTTCTCCCTGATGGGAAA 539
Db |||||||
QY 141 AspGlyAspTyrGluAspGlyThrSerArgSerGlnHisLeuArgPheProAspArg 160
QY 540 CCTTTTCCTCACCAACCCCGGATGGAGAAAGATGGAATTTTCATCTACGTTCTTCCACACATT 599
Db |||||||
QY 161 ProPheProArgProHisGlyTyrLysLysTyrSerPheValTyrValPheHisThrLeu 180
QY 600 GGTCAGTATTTCCAGAAATTTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db |||||||
QY 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerAlaArgValSerIleAsnThrValAsn 200
QY 660 GTGACACTTGGGCTCAACTCATGGAAGTGAAGTGTCTACAGAAAGACATGGACGGGCATAT 719
Db |||||||
QY 201 LeuThrAlaGlyProGlnValMetGluValThrValPheArgArgTyrGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTGAACAGATCAGATTCCTGTGTTTGTG 779
Db |||||||
QY 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db |||||||
QY 241 ThrMetSerGlnLysAsnAspArgAsnLeuSerAspGluIlePheLeuArgAspLeuPro 260
QY 839 ATTATGTTTGTATGCTCTGATTCATGATTCCTAGCCACTTCTCAATTATTCACCATTAAC 898
Db |||||||
QY 261 IleValPheAspValLeuIleHisAspProSerHisPheLeuAsnAspSerAlaIleSer 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCCTGTTTGTTCACCAATCATATCTGTGAAT 958
Db |||||||
QY 281 TyrLysTyrAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCCTTAACCTCACTGTGAAAGCTGCAGCACCA 1018
Db |||||||
QY 301 HisThrTyrValLeuAsnGlyThrPheAsnLeuAsnLeuThrValGlnThrAlaValPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCA-----CCACCCAGACCT----- 1054
Db |||||||
QY 321 GlyProCysProProProSerProSerThrProProSerProSerThrProProLeuPro 340
QY 1055 -----TCAAAACCCACCCCTTCTTTAGGACCTGCTGCTGTGAC 1090
Db |||||||
QY 341 SerProSerProLeuProThrLeuSerThrProSerProSerLeuMetProThrGlyTyr 360
QY 1091 AACCCCTTGAGCTGAGTAGGATTCCTGATGAAAACCTGCCAGATTAACAGATATGGCCAC 1150
Db |||||||
QY 361 LysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyr 380
QY 1151 TTCAAGCCACCATCACAAATGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACA 1210
Db |||||||
QY 381 PheArgAlaThrIleThrIleValGluGlyIleLeuGluValSerIleMetGlnIleAla 400
QY 1211 GACGTCCTGATGCCGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGC 1270
Db |||||||
QY 401 AspValProMetProThrProGlnProAlaAsnSerLeuMetAspPheThrValThrCys 420
QY 1271 CAAGGAGCATTCCTCCAGGAGTCTGTACCATCATTTCTGACCCACCTGCGAGATCACC 1330
Db |||||||
QY 421 LysGlyAlaThrProMetGluAlaCysThrIleIleSerAspProThrCysGlnIleAla 440
QY 1331 CAGAACACAGTCTGCAGCCCTGTGGATGTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGA 1390
Db |||||||
QY 441 GlnAsnArgValCysSerProValAlaValAspGlyLeuCysLeuLeuSerValArgArg 460
QY 1391 ACCTTCAATGGTCTGGGACGCTACTGTGTGTAACCTCAACCTGGGGGATGACACAAGCCTG 1450
Db |||||||
QY 461 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeu 480
QY 1451 GCTCTCAGGAGCACCCCTGATTTCTGTCTCTGACAGAGACCCAGCCTCGCCTTTAAGGATG 1510
Db |||||||
```

Db	481	AlaLeuThrSerThrLeuIleSerIleProGlyLysAspProAspSerProLeuArgAla	500
QY	1511	GCAAAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTGTCACTGTGATCTCCCTC	1570
Db	501	ValAsnGlyValLeuIleSerIleGlyCysLeuAlaValLeuValThrMetValThrIle	520
QY	1571	TTGGTGTAACAAAAACAAGGAATACAACCAATAGTCCCTGGGAATGTGGTC	1630
Db	521	LeuLeuTyrLysLysHisLysAlaTyrLysProIleGlyAsnCysProArgAsnThrVal	540
QY	1631	AGAAGCAAAAGGCCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGAAAC	1690
Db	541	LysGlyLysGlyLeuSerValLeuLeuSerHisAlaLysAlaProphePheArgGlyAsp	560
QY	1691	CAGGAAAGGATCCGCTACTCAAAAACCAA	1720
Db	561	GlnGluLysAspProLeuLeuGlnAspLys	570
RESULT 4			
Q8BVV9	PRELIMINARY; PRT; 574 AA.		
ID	Q8BVV9		
AC	Q8BVV9;		
DT	01-MAR-2003 (TrEMBLrel. 23, Created)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Glycoprotein.		
GN	GNMB.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Skin;		
RX	MEDLINE=22354683; PubMed=12466851;		
RA	The FANTOM Consortium,		
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;		
RT	"Analysis of the mouse transcriptome based on functional annotation of		
RT	60,770 full-length cDNAs.";		
RL	Nature 420:563-573 (2002).		
DR	EMBL; AK076347; BAC36306.1; -.		
DR	MGD; MGI:1934765; Gpnm.		
DR	GO; GO:0005887; C:integral to plasma membrane; IDA.		
DR	GO; GO:0005194; F:cell adhesion molecule activity; IDA.		
DR	GO; GO:0008201; F:heparin binding; IDA.		
DR	GO; GO:0005178; F:integrin binding; IDA.		
DR	GO; GO:0007155; P:cell adhesion; IDA.		
DR	InterPro; IPR000583; GATase_2.		
DR	InterPro; IPR00601; PKD.		
DR	Pfam; PF00801; PKD; 1.		
DR	SMART; SM00089; PKD; 1.		
DR	PROSITE; PS00443; GATASE_TYPE_II; 1.		
DR	PROSITE; PS50093; PKD; 1.		
SQ	SEQUENCE 574 AA; 63675 MW; B6AE99DD5DB7C6D0 CRC64;		
Alignment Scores:			
Pred. No.:	3.05e-202	Length:	574
Score:	2170.00	Matches:	402
Percent Similarity:	82.63%	Conservative:	69
Best Local Similarity:	70.53%	Mismatches:	83
Query Match:	44.74%	Indels:	17
DB:	11	Gaps:	2
US-10-039-272-1 (1-2661) x Q8BVV9 (1-574)			
QY	60	ATGGAATGCTCTACTATTCTCTGGGATTCTGCTCTGGTGCAAGATTGCCACTTGAT	119
Db	1	MetGluSerLeuCysGlyValLeuGlyPheLeuLeuAlaAlaGlyLeuProLeuGln	20
QY	120	GCCGCCAAACGATTTCATGATGTGCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAG	179
Db	21	AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu	40

QY	180	CACAATCAATTAATGGCTGCTTCTGATGAAAAATGACTGGAATGAAAAACTTACCCA	239
Db	41	HisAsnGlnLeuAargGlyTrpSerSerAspGluAsnGluTrpAspGluHisLeuTyrPro	60
QY	240	GTGTGGAACGGGGAGACATGAGGTGGAATAAACTCCTGGAAGGGAGCGCTGTCAGGCG	299
Db	61	ValTrpAargGlyAspGlyAargTrpLysAspSerTrpGluGlyGlyArgValGlnAla	80
QY	300	GTCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTTGCCGTGAACCTG	359
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu	100
QY	360	ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC	419
Db	101	ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120
QY	420	AGAAATGAGGCTGGTTTATCTGCTGATCCATATATGTTTACAACCTGGACAGCATGGTCAGAG	479
Db	121	ArgAsnAspLeuGlyLeuThrSerAspLeuHisValTyrAsnTrpThrAlaGlyAlaAsp	140
QY	480	GACAGTGACGGGAAAATGGCACCGGCCAAAGCCATCAFAACGTCTTCCCTGATGGGAAA	539
Db	141	AspGlyAspTrpGluAspGlyThrSerArgSerGlnHisLeuAargPheProAspAargArg	160
QY	540	CCTTTTCCCTCACCAACCCCGGATGGAGAAGATGGAATTTCATCTACGTCTTCCACACACTT	599
Db	161	ProPheProArgProHisGlyTrpLysLysTrpSerPheValTyrValPheHisThrLeu	180
QY	600	GCTCAGTATTTCCAGAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACACAGCCAAT	659
Db	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerAlaAargValSerIleAsnThrValAsn	200
QY	660	GTGACACTTGGGCCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATAT	719
Db	201	LeuThrAlaGlyProGlnValMetGluValThrValPheAargTyrGlyAargAlaTyr	220
QY	720	GTTCCCATCGCACAAAGTGAAGATGTGTACGTGTAACAGATCAGATTCTGTGTTGTG	779
Db	221	IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProValPheVal	240
QY	780	ACTATGTTCCAGAAACGATCGAAATTTCATCCGACGAAACCTTCCC-AAAGATCTCCCC	838
Db	241	ThrMetSerGlnLysAsnAspArgAsnLeuSerAspGluIlePheLeuAargAspLeuPro	260
QY	839	ATTATGTTTGATGTCCTGATTTCATGATCCTAGCCACTTCTCTCAATTATTCTACCATTAAAC	898
Db	261	IleValPheAspValLeuIleHisAspProSerHisPheLeuAsnAspSerAlaIleSer	280
QY	899	TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATCTGTGAAT	958
Db	281	TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn	300
QY	959	CACACGTATGTCTCAATGGAACCTTCAGCCCTTAACCTCACTGTGAAAGCTGCAGCACCA	1018
Db	301	HisThrTyrValLeuAsnGlyThrPheAsnLeuAsnLeuThrValGlnThrAlaValPro	320
QY	1019	GGACCTTGTCGG-----TCAAAAACCCACCCCTTCTTTAGGACCTGCTGGTGAC	1042
Db	321	GlyProCysProProProSerProSerThrProProProProSerThrProProSerPro	340
QY	1043	CCACCCAGACCT-----TCAAAAACCCACCCCTTCTTTAGGACCTGCTGGTGAC	1090
Db	341	ProProSerProLeuProThrLeuSerThrProSerProSerProSerLeuMetProThrGlyTyr	360
QY	1091	AACCCCTGGAGCTGAGTAGGATTCTCTGATGAAAACTGCCAGATTAAACAGATATGGCCAC	1150
Db	361	LysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnAargTyrGlyTyr	380
QY	1151	TTTCAAGCCACCATCACAAATTGTAGAGGGAATCTTAGAGGTTAACATCATCCAGATGACA	1210
Db	381	PheArgAlaThrIleThrIleValGluGlyIleLeuGluValSerIleMetGlnIleAla	400

```
QY 1211 GACGTCCTGATGCCGGTGCCTGAAAGCTCCCTAATAGACTTTGTCTGACCTGC 1270
    ||||| ||||| ||| ||| :|||:||||| ||||| |||||
Db 401 AspValProMetProThrProGlnProAlaAsnSerLeuMetAspPheThrValThrCys 420
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1271 CAAGGAGCATTCCCAAGGAGTCTGTACCATCATTTCTGACCCACCTGCGAGATCAC 1330
    :|||: ||| ||||| ||||| ||||| ||||| ||||| ||||| :|||:
Db 421 LysGlyAlaThrProMetGluAlaCysThrIleIleSerAspProThrCysGlnIleAla 440
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1331 CAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCTGACGTGAGACGA 1390
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 441 GlnAsnArgValCysSerProValAlaValAspGlyLeuCysLeuLeuSerValArgArg 460
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1391 ACCTTCAATGGTCTGGACGTACTGTGTGAACCTCACCTGGGGGATGACACAAGCCTG 1450
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 461 AlaPheAsnGlySerGlyThrTyrcysValAsnPheThrLeuGlyAspAspAlaSerLeu 480
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1451 GCTCTCACGAGCACCTGATTTCTGTCTCTGACAGACCCAGCCTGCGCTTTAAGGATG 1510
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 481 AlaLeuThrSerThrLeuIleSerIleProGlyLysAspProAspSerProLeuArgAla 500
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1511 GCAACAGTGCCTGATCTCCGTTGGCTGTGGCCATATTTGTCACTGTGATCTCCCTC 1570
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 501 valAsnGlyvalleuIleSerIleGlyCysLeuAlaValLeuValThrMetValThrIle 520
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1571 TTGGTGTACAAAAACACAAGGAATACACCCCAATAGAAATAGTCTCTGGGAATGTGTC 1630
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 521 LeuLeuTyrlLysLysHisLysAlaTyrlLysProIleGlyAsnCysProArgAsnThrVal 540
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1631 AGAAGCAAGGCCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTCCGGGAAAC 1690
    :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| :|||:
Db 541 LysGlyLysGlyLeuSerValLeuLeuSerHisAlaLysAlaProPhePheArgGlyAsp 560
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1691 CAGGAAAAGGATCCGCTACTCAAAAACCAA 1720
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 561 GlnGluLysAspProLeuLeuGlnAspLys 570
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

RESULT 5

```
Q8BXL4
ID Q8BXL4 PRELIMINARY; PRT; 574 AA.
AC Q8BXL4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glycoprotein.
GN GPNMB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Retina;
RC MEDLINE=22354683; PubMed=12466851;
RX The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK044764; BAC32074.1; -.
DR MGD; MGI:1934765; Gpmb.
DR GO; GO:0005887; C:integral to plasma membrane; IDA.
DR GO; GO:0005194; F:cell adhesion molecule activity; IDA.
DR GO; GO:0008201; F:heparin binding; IDA.
DR GO; GO:0005178; F:integrin binding; IDA.
DR GO; GO:0007155; P:cell adhesion; IDA.
DR InterPro; IPR000583; GATase_2.
DR InterPro; IPR000601; PKD.
DR Pfam; PF00801; PKD; 1.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS00443; GATASE_TYPE_II; 1.
DR PROSITE; PS50093; PKD; 1.
SQ SEQUENCE 574 AA; 63693 MW; DC0D48485BA2E9D1 CRC64;
```

```
Alignment Scores:
Pred. No.: 7.5e-202 Length: 574
Score: 2166.00 Matches: 401
Percent Similarity: 82.46% Conservative: 69
Best Local Similarity: 70.35% Mismatches: 84
Query Match: 44.66% Indels: 17
DB: 11 Gaps: 2

US-10-039-272-1 (1-2661) x Q8BXL4 (1-574)

QY 60 ATGGAATGTCTCTACTATTTCCTGGGATTTCTGCTCTGGCTGCAAGATTGCCACTTGAT 119
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 MetGluSerLeuCysGlyValLeuGlyPheLeuLeuAlaGlyLeuProLeuGln 20
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 120 GCCGCCAAACGATTTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrrProAsnHisMetArgGlu 40
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 180 CACAATCAATTAAATGGCTGGTCTTCTGATGAAAAATGACTGGAATGAAAAAACTCTACCCA 239
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 41 HisAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGluHisLeuTyrrPro 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 240 GTGTGGAAAGCGGGGAGACATGAGGTGGAATAAACTCCTGGAAGGGAGGCCGTGTGCAGGCG 299
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 ValTrpArgArgGlyAspGlyArgTrpLysAspSerTrpGluGlyArgValGlnAla 80
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 300 GTCCTGACCACTGACTCACCGCCCTCGTGGCTCAAATATAACATTTGCGGTGAACCTG 359
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 360 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
    :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| :|||:
Db 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrrGluLysAsnCys 120
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 420 AGAAATGAGGCTGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG 479
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 ArgAsnAspLeuGlyLeuThrSerAspLeuHisValTyrrAsnTrpThrAlaGlyAlaAsp 140
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 480 GACAGTGACGGGGAATGGCACCGGCCCAAGCCATCAACGTTCTCCCTGATGGGAAA 539
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 141 AspGlyAspTrpGluAspGlyThrSerArgSerGlnHisLeuArgPheProAspArgArg 160
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 540 CCTTTTCTCACCCCGGATGGAGAAGATGGGAATTTTCATCTACGTTCTCCACACACTT 599
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 161 PropheProArgProHisGlyTrpLysLysTrpSerPheValTyrrPheHisThrLeu 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 600 GGTCAAGTATTTCCAGAAATGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACACGCCAAT 659
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 GlyGlnTyrrPheGlnLysLeuGlyArgCysSerAlaArgValSerIleAsnThrValAsn 200
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 660 GTGACACTTGGGCCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATAT 719
    :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| :|||:
Db 201 LeuThrAlaGlyProGlnValMetGluValThrValPheArgArgTyrrGlyArgAlaTyrr 220
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCTGTGTTTGTG 779
    :|||: ||||| ||||| ||||| ||||| ||||| ||||| ||||| :|||:
Db 221 IleProIleSerLysValLysAspValTyrrValIleThrAspGlnIleProValPheVal 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 780 ACTATGTTCCAGAGAAGACGATCGAAATTCATCCGACGCAAAACCTTCCC-AAAGATCTCCCC 838
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 ThrMetSerGlnLysAsnAspArgAsnLeuSerAspGluIlePheLeuArgAspLeuPro 260
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 839 ATTATGTTTGTATGTCCTGATTATCATGATCCTAGCCACTTCCCTCAATATTCTACCATTAAC 898
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 261 IleValPheAspValLeuIleHisAspProSerHisPheLeuAsnAspSerAlaIleSer 280
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTGTTTCCACCAATCATACTGTGAAT 958
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 281 TyrrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 959 CACACGATGTGCTCAATGGAACCTTCAGCCTTAACCTCAGCTGTGAAAGCTGCAGCACCA 1018
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 HisThrTyrrValLeuAsnGlyThrPheAsnLeuAsnLeuThrValGlnThrAlaValPro 320
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


QY 1019 GGACCTTGTCG-----CCACCGCCACCA 1042
Db 321 GlyProCysProProSerProSerThrProProProSerThrProProSerPro 340
QY 1043 CCACCCAGACCT-----TCAAAACCCACCCCTTCTTTAGGACCTGCTGGTGAC 1090
Db 341 ProProSerProLeuProThrLeuSerThrProSerProSerLeuMetProThrGlyTyr 360
QY 1091 AACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACTGCCAGATTAAACAGATATGGCCAC 1150
Db 361 LysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyr 380
QY 1151 TTTCAAGCCACCATCACAAATTGTAGAGGAATCTTAGAGGTTAAACATCATCCAGATGACA 1210
Db 381 PheArgAlaThrIleThrIleValGluGlyIleLeuGluValSerIleMetGlnIleAla 400
QY 1211 GACGTCCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGC 1270
Db 401 AspValProMetProThrProGlnProAlaAsnSerLeuMetAspPheThrValThrCys 420
QY 1271 CAAGGAGCATTCGCCAGGCTGTGTACCATCATTTCTGACCCACCTGCGAGATCAC 1330
Db 421 LysGlyAlaThrProMetGluAlaCysThrIleIleSerAspProThrCysGlnIleAla 440
QY 1331 CAGAACACAGCTGCGAGCCCTGTGGATGTGGATGAGATGTCGTGCTGACTGTGAGACGA 1390
Db 441 GlnAsnArgValCysSerProValAlaValAspGlyLeuCysLeuSerValArgArg 460
QY 1391 ACCTTCAATGGTCTGGGACGFACTGTGTGAACCTCACCTGGGGATGACACAAGCCTG 1450
Db 461 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeu 480
QY 1451 GCTCTCACGAGCACCCCTGATTTCTGTTCTCTGACAGAGACCCAGCCTGCGCTTTAAGGATG 1510
Db 481 AlaLeuThrSerThrLeuIleSerIleProGlyLysAspProAspSerProMetArgAla 500
QY 1511 GCAAAACAGTGCCTGATCTCCGTGGCTGGCTGTGGCCATATTGTACATGTCATCCTCCTC 1570
Db 501 ValAsnGlyValLeuIleSerIleGlyCysLeuAlaValLeuValThrMetValThrIle 520
QY 1571 TTGGTGTAACAAAAACAAGGAATACACCCCAATAGAAAAATAGTCTGGGAATGTGGTC 1630
Db 521 LeuLeuTyrLysLysHisLysAlaTyrLysProIleGlyAsnCysProArgAsnThrVal 540
QY 1631 AGAAGCAAAGCCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGAAAC 1690
Db 541 GluGlyLysGlyLeuSerValLeuLeuSerHisAlaLysAlaProPhePheArgGlyAsp 560
QY 1691 CAGGAAAAGGATCCGCTACTCAAAAACCAA 1720
Db 561 GlnGluLysAspProLeuLeuGlnAspLys 570

RESULT 6

Q9QZF6 PRELIMINARY; PRT; 572 AA.
AC Q9QZF6
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Osteoactivin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu J., Safadi F.F., Smock S.L., Rosenzweig A.B., Odgren P.R.,
RA Marks S.C. Jr., Popoff S.N., Owen T.A.,
RT "Cloning and characterization of a novel cDNA highly expressed in
RT osteopetrotic bone."
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF184983; AAF03400.1; -.

DR InterPro; IPR000601; PKD.
DR SMART; SM00089; PKD; 1.
DR PROSITE; PS50093; PKD; 1.
SQ SEQUENCE 572 AA; 63772 MW; CB939FB043EC57EB CRC64;
Alignment Scores:
Pred. No.: 6.33e-199 Length: 572
Score: 2136.00 Matches: 393
Percent Similarity: 82.39% Conservative: 75
Best Local Similarity: 69.19% Mismatches: 86
Query Match: 44.04% Indels: 15
DB: 11 Gaps: 1

US-10-039-272-1 (1-2661) x Q9QZF6 (1-572)

QY 60 ATGGAATGTCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuAlaAlaGlyLeuProLeuGln 20
QY 120 GCCGCCAAACGATTTCATGATGTGCTGGCAATGAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheArgAspValLeuGlyHisGlnGlnTyrProAspHisMetArgGlu 40
QY 180 CACAATCAATTAAATGCTGCTCTTCTGATGAAATCACTGGAATGAAAACTCTACCCA 239
Db 41 AsnAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGluGlnLeuTyrPro 60
QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAAAAAATCTGGAAGGGAGGCCGTGTCAGGCG 299
Db 61 ValTrpArgArgGlyGluGlyArgTrpLysAspSerTrpGluGlyArgValGlnAla 80
QY 300 GTCCTGACCAGTCACTCACCCAGCCCTCGTGGGCTCAAAATATAACATTTGGCGTGAACCTG 359
Db 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120
QY 420 AGAAATGAGGCTGTTTATCTGCTGATCCATATGTTTCAACTGGACAGCATGGTCAGAG 479
Db 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAsp 140
QY 480 GACAGTGACGGGAAAATGGCACCGGCCAAAGCCATCAACGCTCTCCCTGATGGGAAA 539
Db 141 AspGluAspTrpGluAspAsnThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160
QY 540 CTTTTCCTCACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCTTCCACACACTT 599
Db 161 PropheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180
QY 600 GGTCAAGTATTTCCAGAAATTTGGACGATGTTTCAGTGAGAGATTTCTGTGACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyGlnCysSerAlaArgValSerIleAsnThrValAsn 200
QY 660 GTGACACTTGGCCCTCAACTCATGGAAGTGACTGTCTACAGAGACATGGACGGCATAT 719
Db 201 LeuThrValGlyProGlnValMetGluValIleValPheArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAACAGATCAGATTCTGTGTTGTG 779
Db 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240
QY 780 ACTATGTTCCAGAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260
QY 839 ATTATGTTTGATGTCTGATTCATGATCCTAGCCACTTCTCCTCAATTATTCTACCATTAC 898
Db 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATACTGTGAAT 958

RX	MEDLINE=22354683; PubMed=12466851;		
RA	The FANTOM Consortium,		
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;		
RT	"Analysis of the mouse transcriptome based on functional annotation of		
RT	60,770 full-length cDNAs.";		
RL	Nature 420:563-573 (2002).		
DR	EMBL; AK079220; BAC37578.1; ..		
DR	GO; GO:0008152; P:metabolism; IEA.		
DR	InterPro; IPR00583; GATase_2.		
DR	InterPro; IPR00601; PKD.		
DR	Pfam; PF00801; PKD; 1.		
DR	SMART; SM00089; PKD; 1.		
DR	PROSITE; PS00443; GATASE_TYPE_II; 1.		
DR	PROSITE; PS50093; PKD; 1.		
SQ	SEQUENCE 526 AA; 58250 MW; 82B93B49E341EC72 CRC64;		
Alignment Scores:			
Pred. No.:	3.91e-187	Length:	526
Score:	2015.00	Matches:	372
Percent Similarity:	83.17%	Conservative:	63
Best Local Similarity:	71.13%	Mismatches:	72
Query Match:	41.55%	Indels:	17
DB:	11	Gaps:	2
US-10-039-272-1 (1-2661) x Q8BVA0 (1-526)			
QY	60	ATGGAATGTCCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT	119
Db	1	MetGluSerLeuCysGlyValLeuGlyPheLeuLeuAlaGlyLeuProLeuGln	20
QY	120	GCCGCCAAACGATTTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG	179
Db	21	AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu	40
QY	180	CACAATCAATTAAATGGTGGTCTTCTGATGAAATGACTGGAATGAAAACTCTACCCA	239
Db	41	HisAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGluHisLeuTyrPro	60
QY	240	GTGTGGAAGCGGGAGACATGAGGTGGAATACTCCTGGAAGGGAGGCCGTGTGCAGGCG	299
Db	61	ValTrpArgArgGlyAspGlyArgTrpLysAspSerTrpGluGlyArgValGlnAla	80
QY	300	GTCTGACCAAGTACTCACAGCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTG	359
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu	100
QY	360	ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC	419
Db	101	ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysasnCys	120
QY	420	AGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGGTCAGAG	479
Db	121	ArgAsnAspLeuGlyLeuThrSerAspLeuHisValTyrAsnTrpThrAlaGlyAlaAsp	140
QY	480	GACAGTGACGGGAAAATGGCACCGCCAAAGCCATCATACCGTCTTCCCTGATGGGAAA	539
Db	141	AspGlyAspTrpGluAspGlyThrSerArgSerGlnHisLeuArgPheProAspArgArg	160
QY	540	CCTTTTCTCACCCCGGATGGAGAGATGGAATTTTCATCTACGTCTTCCACACACTT	599
Db	161	ProPheProArgProHisGlyTrpLysLysTrpSerPheValTyrValPheHisThrLeu	180
QY	600	GGTCAGTATTTCCAGAAATTCGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT	659
Db	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerAlaArgValSerIleAsnThrValAsn	200
QY	660	GTGACACTTGGGCTCAACTCATGGAAGTACTGTCTACAGAAGACATGGACGGGCATAT	719
Db	201	LeuThrAlaGlyProGlnValMetGluValThrValPheArgArgTyrGlyArgAlaTyr	220
QY	720	GTTCCCATCGCACAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCTGTGTTTGTG	779
Db	221	IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProValPheVal	240

```
QY 780 ACTATGTTCCAGAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetSerGlnLysAsnAspArgAsnLeuSerAspGluIlePheLeuArgAspLeuPro 260
QY 839 ATTATGTTTGATGTCTGATTCATGATCCTAGCCACTTCCCTCAATTATTCTACCATTAAC 898
Db 261 IlevAlPheAspValLeuIleHisAspProSerHisPheLeuAsnAspSerAlaIleSer 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATACTGTGAAT 958
Db 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
QY 959 CACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheAsnLeuAsnLeuThrValGlnThrAlaValPro 320
QY 1019 GGACCTTGTCGG-----CCACCGCCACCA 1042
Db 321 GlyProCysProProProSerProSerThrProProProSerThrProProSerPro 340
QY 1043 CCACCCAGACCT-----TCAAAAACCCACCCCTTCTTTAGGACCTGCTGGTGAC 1090
Db 341 ProProSerProLeuProThrLeuSerThrProSerProSerLeuMetProThrGlyTyr 360
QY 1091 AACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTAAACAGATATGGCCAC 1150
Db 361 LysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyr 380
QY 1151 TTTCAGCCACCATCACAAATTGTAGAGGAATCTTAGAGGTTAAACATCATCCAGATGACA 1210
Db 381 PheArgAlaThrIleThrIleValGluGlyIleLeuGluValSerIleMetGlnIleAla 400
QY 1211 GACGTCCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGC 1270
Db 401 AspValProMetProThrProGlnProAlaAsnSerLeuMetAspPheThrValThrCys 420
QY 1271 CAAGGGAGCATTCACCGGAGGTCTGTACCATCATTTCTGACCCACCTGGAGATCACC 1330
Db 421 LysGlyAlaThrProMetGluAlaCysThrIleIleSerAspProThrCysGlnIleAla 440
QY 1331 CAGAACACAGCTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGA 1390
Db 441 GlnAsnArgValCysSerProValAlaValAspGlyLeuCysLeuLeuSerValArgArg 460
QY 1391 ACCTTCAATGGGTCTGGAGCTACTGTGTGAACCTCACCCCTGGGGATGACACAAGCCTG 1450
Db 461 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeu 480
QY 1451 GCTCTCAGCAGCACCTGATTTCTGTTCTCTGACAGAGACCCAGCCTCGCCTTAAAGGATG 1510
Db 481 AlaLeuThrSerThrLeuIleSerIleProGlyLysAspProAspSerProLeuArgAla 500
QY 1511 GCAAACAGTGCCCTGATCTCGTTGGCTGCTTGGSCATATTGTGCTGATCTCCCTC 1570
Db 501 ValAsnGlyValLeuIleSerIleGlyCysLeuAlaValLeuValThrMetValThrIle 520
QY 1571 TTGGTGATC 1579
Db 521 LeuLeuTyr 523

RESULT 8
Q96F58 PRELIMINARY; PRT; 206 AA.
AC Q96F58;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to glycoprotein (Transmembrane) nmb (Hypothetical protein
DE GPNMB).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99063792; PubMed=9847074;
RA Sulston J.E., Waterston R.;
RT "Toward a complete human genome sequence.";
RL Genome Res. 8:1097-1108(1998).
RN [3]
RP SEQUENCE FROM N.A.
RA Kalicki J., Kramer J.;
RT "The sequence of Homo sapiens BAC clone CTA-271G13.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA Waterston R.H.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
RN [8]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
RN [9]
RP SEQUENCE FROM N.A.
RA Kalnine N., Chen X., Rolfs A., Halleck A., Hines L., Eisenstein S.,
RA Koundinya M., Raphael J., Moreira D., Kelley T., LaBaer J., Lin Y.,
RA Phelan M., Farmer A.;
RT "Cloning of human full-length CDSS in BD Creator(TM) System Donor
RT vector.";
RL Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR EMBL; BC011595; AAH11595.1; -.
DR EMBL; AC005082; AAP22337.1; -.
DR EMBL; BT007074; AAP35737.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
KW Hypothetical protein; Transmembrane.
SQ SEQUENCE 206 AA; 23945 MW; 30CBDE6928D73FBD CRC64;

Alignment Scores:
Pred. No.: 4.26e-91 Length: 206
Score: 1029.00 Matches: 182
Percent Similarity: 98.38% Conservative: 0
Best Local Similarity: 98.38% Mismatches: 3
Query Match: 21.22% Indels: 0
DB: 4 Gaps: 0

US-10-039-272-1 (1-2661) x Q96F58 (1-206)
QY 60 ATGGAATGCTCTACTATTCTCTGGGATTTCTGCTCTGGTGCAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrThrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAATGGCTGCTCTTCTGATGAAATGACTGGAATGAAAAACTCTACCCA 239
```


SQ	SEQUENCE	626 AA;	66301 MW;	7EC0A06C63212674	CRG64;	
Alignment Scores:						
Pred. No.:	2.18e-53	Length:	626			
Score:	644.50	Matches:	185			
Percent Similarity:	40.94%	Conservative:	93			
Best Local Similarity:	27.25%	Mismatches:	210			
Query Match:	13.29%	Indels:	192			
DB:	11	Gaps:	21			
US-10-039-272-1 (1-2661) x Q9CZB2 (1-626)						
QY	78	TTCTGGGATTTC	TGCTCCTGGCTGCAAGATTGCCACTTGATGCC-----GCCAAACGA	131		
Db	8	PheLeuProValLeuValLeuSerAlaLeuLeuAlaValaiglyAlaLeuGluGlySerArg	27			
QY	132	TTTCATGATGTG	TGGGCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTA	191		
Db	28	AsnGlnAspTrpLeuGlyValProArgGlnLeuVal-----	39			
QY	192	AATGGCTGGTCT	TCTGATGAAAAATGACTGGAATGAAAAAATCTACCCAGTGTGGAAACGG	251		
Db	40	-----	ThrLysThrTrpAsnArgGlnLeuTyrProGluTrpThr---	52		
QY	252	GGAGACATGAGGT	GTGAAAAAATCCTGGAAGGAGGCGGTGTGCAGGCGTCTCGACCAAGT	311		
Db	53	---GluValGlnGlySerAsnCysTrpArgGlyGlyGlnValSerLeuArgValIleAsn	71			
QY	312	GACTCACCAAGCA	ATGTCGTTGGCTCAAATATAACATTTTCGGGTGAACCTGATATTCCTTAGA	371		
Db	72	AspGlyProThrLeuValGlyAlaAsnAlaSerPheSerIleAlaLeuHisPheProGly	91			
QY	372	TGCCAAAAGGAAG	ATGCCAATGGCAACATAGTCTATGAGAAGAAC---TGCAGAAATGAG	428		
Db	92	SerGlnLysValLeuProAspGlyGlnValIleTrpAlaAsnAsnThrIleIleAsnGly	111			
QY	429	GCTGTTTATCTG	CTATCCATATGTTTACAACCTGGACAGCATGGTCAGAGGACAGTGAC	488		
Db	112	SerGlnValTrpGlyGlnProValTyrPro-----	GlnGluProAspAsp	127		
QY	489	GGGAAAAATGGC	ACCGGCCAAAAGCCATCATACGCTTCCCTGATGGGAAACCTTTTCCT	548		
Db	128	Ala-----	CysValPheProAspGlyGlyProCysPro	138		
QY	549	CACCACCCCGGAT	GGAGAAGATGGAATTTCATCTACGTCTTCCACACACACTTGGTCAGTAT	608		
Db	139	SerGlyProLysProProLysArgSerPheValTyrValTrpLysThrTrpGlyLysTyr	158			
QY	609	TTCCAGAAATTGG	ACCATGTTTCAGTGAGAGTTTCTGTGTAACACACAGCCCAATGTGACACTT	668		
Db	159	TrpGlnValLeuGlyGlyProValSerArgLeuSerIleAlaThrGlyHisAlaLysLeu	178			
QY	669	GGGCCTCAACTC	ATGGAAGTGACTGTCTACAGAAGACATGGA---CGGGCATATGTTCCC	725		
Db	179	GlyThrHisThrMetGluValThrValTyrHisArgArgGlySerGlnSerTyrValPro	198			
QY	726	ATCGCACAAAGT	GAAAGATGTGTACGTGGTTAACAGATCAGATTCTCTGTGTTGTGACTATG	785		
Db	199	LeuAlaHisAlaSerSerThrPheThrIleThrAspGlnValProPheSerValSerVal	218			
QY	786	TTCCAGAA-----	GAACGATCGAAATTCATCCGACGCAAAACCTTCCCAA	829		
Db	219	SerGlnLeuGlnAlaLeuAspGlyGluThrLysHisPheLeuArgAsnHis-----	235			
QY	830	GATCTCCCCATTA	TGTTTGATGTCCTGATTCATGATCCATGACCTTCCTCAATATTCT	889		
Db	236	-----ProLeuIlePheAlaLeuGlnLeuHisAspProSerGlyTyrLeuAlaGluAla	253			
QY	890	ACCATTAACTACA	AGTGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCAT	949		
Db	254	AspLeuSerTyrThrTrpAspPheGlyAspGlyThrGlyThrLeuIleSerArgAlaLeu	273			
QY	950	ACTGTGAATCAC	ACGTATGTGCTCAATGGAAACCTTCAGCCTTAACCTCACTGTGAAAGCT	1009		

Db	274	AspValThrHisThrTyrLeuGluSerGlySerValThrAlaGlnValValLeuGlnAla	293			
QY	1010	GCA-----				
Db	294	AlaIleProLeuValSerCysGlySerSerProValProGlyThrThrAspGlyTyrMet	313			
QY	1012	-----				
Db	314	ProThrAlaGluAlaProGlyThrThrSerArgGlnGlyThrThrLysValValGly	333			
QY	1013	---GCACCAGGAC	CTTGTCCGCCACCGCCACCA-----	1042		
Db	334	ThrThrProGlyGlnMetProThrThrGlnProSerGlyThrThrValValGlnMetPro	353			
QY	1042	-----				
Db	354	ThrThrGluValThrAlaThrThrSerGluGlnMetLeuThrSerAlaValIleAspThr	373			
QY	1043	-----	CCACCCAGACCTTCA---	1057		
Db	374	ThrLeuAlaGluValSerThrThrGluGlyThrGlyThrThrProThrArgProSerGly	393			
QY	1058	-----AAACCCAC	CCCTTCTTTAGGACCTGCTGGTGACAACCCCTG-----	1099		
Db	394	ThrThrValAlaGlnAlaThrThrThrGluGlyProAspAla---SerProLeuLeuPro	412			
QY	1100	-----GAGCTG	AGTAGGATTCTCTGATGAA-----	1123		
Db	413	ThrGlnSerSerThrGlySerIleSerProLeuLeuAspAspThrAspThrIleMetLeu	432			
QY	1124	-----AACTGCC	AGATTAAACAGATATGGCCACTTTTCAAGCCACC	1162		
Db	433	ValLysArgGlnValProLeuAspCysValLeuTyrArgTyrGlySerPheSerLeuAla	452			
QY	1163	ATCACAAATTGT	AGGGGAATCTTAGAGGTTTAACATCATCCAGATGACAGACGCTCCTGATG	1222		
Db	453	LeuAspIleValGlnGly-----	IleGluSerAlaGluIleLeuGln	466		
QY	1223	CCGGTGCCATG	GCCCTGAAAGCTCCCTAATAAGACTTTGTGCTGACCTGCCAAGGGAGCATT	1282		
Db	467	AlaValProPheSerGluGlyAspAlaPheGluLeuThrValSerCysGlnGlyGlyLeu	486			
QY	1283	CCCACGGAGGT	CTGTACCATCATTTCTGACCCACCTCGAGATCACCAGAACACAGTC	1342		
Db	487	ProLysGluAlaCysMetAspIleSerSerProGlyCysGlnProProAlaGlnArgLeu	506			
QY	1343	TGCAGCCCTGT	GATGTGGATGTGATGTGCTGCTGACTGTGAGACGAACCTTC---AAT	1399		
Db	507	CysGlnSerValProProSerProAspCysGlnLeuValLeuHisGlnValLeuLysGly	526			
QY	1400	GGGTCTGGGAC	GCTACTGTGTGAACCTCACCCCTGGGGATGACACAAAGCCTGGCTCTCAGC	1459		
Db	527	GlySerGlyThrTyrCysLeuAsnValSerLeuAlaAspAlaAsnSerLeuAlaValAla	546			
QY	1460	AGCACCCCTGA	TTTCTTCCTGCACAGACCCAGCCCTCGCCTTTAAGGATGGCAAACAGT	1519		
Db	547	SerThrGlnLeuValValProGlyGlnAspGlyGly-----	LeuGlyGln	561		
QY	1520	GCCCTGATCTC	CGTTGGCTGGTTCGCCATATTTGTCACTGTGATCTCCCTCTTGGTGATC	1579		
Db	562	AlaProLeuLeuValGlyIleLeuLeuValLeuValAlaValValLeuAlaSerLeuIle	581			
QY	1580	AAAAACAC-----	AAGGAATACAAACCAATAGAAAAATAGTCTGGGAAT-----	1624		
Db	582	HisArgHisArgLeuLysLysGlnGlySerValSerGlnMetProHisGlySerThrHis	601			
QY	1625	-----GTGGT	CAGAAAGCAAAGGCCTGAGTGTCTTTCTCAACCGTGCA	1666		
Db	602	TrpLeuArgLeuProProValPheArgAlaArgGlyLeuGly-----	615			
QY	1667	AAAGCCGTGTT	CTTCCCGGGAACACGAGGAAAGGATCCCGCTACTCAAAACCAAGAA	1723		

Db 616 -----GluAsnSerProLeuLeuSerGlyGlnGln 625

RESULT 11

O97884

ID O97884 PRELIMINARY; PRT; 461 AA.

AC O97884;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Melanocyte protein 17 (Fragment).

GN PMEL17.

OS Equus caballus (Horse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.

OX NCBI_TaxID=9796;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Skin;

RA Rieder S., Checa-Cortes M.L., Joerg H., Stranzinger G.;

RT "An equine sequence homologous to melanocyte protein 17 (PMEL17)

RT mapped to chromosome 6q23.";

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF076780; AAC97108.1; --.

DR InterPro; IPR000601; PKD.

DR Pfam; PF00801; PKD; 1.

DR SMART; SM00089; PKD; 1.

DR PROSITE; PS50093; PKD; 1.

FT NON_TER 1

FT NON_TER 461

SQ SEQUENCE 461 AA; 49334 MW; 12752AF6C1EC373D CRC64;

Alignment Scores:

Pred. No.: 7.45e-27 Length: 461

Score: 372.00 Matches: 101

Percent Similarity: 37.84% Conservative: 64

Best Local Similarity: 23.17% Mismatches: 119

Query Match: 7.67% Indels: 153

DB: 6 Gaps: 9

US-10-039-272-1 (1-2661) x O97884 (1-461)

QY 702 AGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGTACGTGTAACAGAT 761

Db 1 ArgGlySerGlnThrTyrMetProLeuAlaHisSerArgSerAlaPheThrIleThrAsp 20

QY 762 CAGATTCTCTGTTGTGACTATGTTCCAGAGAACGATCGAAATTCATCCGACGAAACC 821

Db 21 GlnValProPheSerValSerValSerGlnLeuGlnAlaLeuAspGlyArgAsnLysHis 40

QY 822 TTCCC-AAAGATCTCCCCATTATGTTGATGTCTCTGATTCATGATCCTAGCCACTTCCTC 880

Db 41 PheLeuLysAsnGlnProLeuThrPheAlaLeuArgLeuHisAspProSerGlyTyrLeu 60

QY 881 AATTATTACCATTAACACAAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTC 940

Db 61 AlaGlyAlaAspLeuSerTyrThrTipAspPheGlyAspSerThrGlyThrLeuIleSer 80

QY 941 ACCAATCATACTGTGAATCACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACT 1000

Db 81 ArgAlaLeuValValThrHisThrTyrLeuGluSerProValThrAlaGlnValVal 100

QY 1001 GTGAAAGCTGCAGCACCA-----GGACCTTGTCGCCACCG----- 1036

Db 101 LeuGlnAlaAlaileProLeuThrSerCysGlySerSerProValProGlyThrThrGly 120

QY 1036 ----- 1036

Db 121 GlyTyrValProThrAlaGluAlaProGlyThrThrAlaGlyGlnValProThrAlaAsp 140

QY 1037 -----CCACCACCACCCAGACCTTCAAAACCCACC----- 1066

Db 141 valValAsnThrThrProGlyGlnValProThrAlaGluProSerArgThrThrAlaVal 160

QY 1067 -----CCTTCTTTAGGACCT 1081

Db 161 GlnValProThrThrGluValIleSerThrThrProValGlnValProThrAlaGluAsp 180

QY 1082 GCTGGTGACAAACCCCTGGAGCTGAGTAGG----- 1111

Db 181 IleGlyThrThrProGluGlnValSerThrProGluSerLeuGlyThrThrLeuAlaGlu 200

QY 1112 -----ATTCCTGATGAA----- 1123

Db 201 MetProLeuGlnArgLeuLysGlyIleProProGluValSerAsnSerArgSerLeuLeu 220

QY 1123 ----- 1123

Db 221 GlyHisSerAlaGlnValThrValGlnSerTrpTrpLysProGlnTrpArgGlyThrGln 240

QY 1123 ----- 1123

Db 241 ProLeuSerLeuArgValGlnMetProAlaHisSerCysTyrArgAsnTyrArgSer 260

QY 1123 ----- 1123

Db 261 GlnSerProLeuLeuAspGlyThrAlaThrLeuPheLeuValLysArgGlnValProLeu 280

QY 1124 AACTGCCAGATTAAACAGATATGGCCACTTTTCAAGCCACCATCACAAATTGTAGAGGGAATC 1183

Db 281 AspCysValLeuTyrArgTyrGlySerPheSerLeuThrLeuAspIleValGlnGly--- 299

QY 1184 TTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGGTGCCATGGCCTGAAAGC 1243

Db 300 -----IleGluSerAlaGluIleLeuGlnAlaValProSerSerGluGly 314

QY 1244 TCCCTAATAGACTTGTGTCGTGACCTGCCAAGGAGCATTTCCACGGAGGTCTGTACCATC 1303

Db 315 AspAlaPheGluLeuThrValSerCysGlnGlyGlyLeuProLysGluAlaCysMetAsp 334

QY 1304 ATTTCTGACCCACCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGAT 1363

Db 335 IleSerSerProGlyCysGlnProProAlaGlnArgLeuCysGlnProValThrProAsn 354

QY 1364 GAGATGTGTGTGCTGACTGTGAGACGAAACCTTC---AATGGGTCTGGGACGTACTGTGTG 1420

Db 355 ProAlaCysGlnLeuValLeuHisGlnValLeuLysSerGlySerGlyThrTyrCysLeu 374

QY 1421 AACCTCACCCCTGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTTCTCCT 1480

Db 375 AsnValSerLeuAlaAspAlaAsnSerLeuAlaMetValSerThrGlnLeuValMetPro 394

QY 1481 GACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCCTGATCTCCGTTGGCTGC 1540

Db 395 GlyGlnGluAlaGly-----LeuGlyGlnAlaProLeuPheValGlyIle 409

QY 1541 TTGGCCATATTTGTCACT---GTGATCTCCCTCTTGGTGTCACAAAAA 1585

Db 410 LeuLeuValLeuMetAlaMetValLeuAlaSerLeuIleTyrArgArg 425

RESULT 12

O93391

ID O93391 PRELIMINARY; PRT; 270 AA.

AC O93391;

DT 01-NOV-1998 (TrEMBLrel. 08, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Melanosomal matrix protein (Fragment).

GN MPM115.

OS Coturnix coturnix (Common quail).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Coturnix.

OX NCBI_TaxID=9091;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=99119493; PubMed=9918878;

RA Oberst C., Hartl M., Weiskirchen R., Bister K.;
RT "Conditional cell transformation by doxycycline-controlled expression
of the MC29 v-myc allele";
RL Virology 253:193-207(1999).
DR EMBL; AF077328; AAD12180.1; -.
DR GO; GO:0005198; F:structural molecule activity; IEA.
KW Matrix protein.
FT NON TER 1
SQ SEQUENCE 270 AA; 26785 MW; 564AA46F0B875B63 CRC64;

Alignment Scores:
Pred. No.: 3.84e-17 Length: 270
Score: 271.50 Matches: 71
Percent Similarity: 48.03% Conservative: 39
Best Local Similarity: 31.00% Mismatches: 94
Query Match: 5.60% Indels: 25
DB: 13 Gaps: 6

US-10-039-272-1 (1-2661) x Q93391 (1-270)

QY 1067 CCTTCTTTAGGACCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAC 1126
|||:|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 50 ProThrAlaGlySerThrAlaGluProLeuLeuLeuValLysArgGlnAlaProGluSer 69
|||:|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1127 -----TGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCACAATTGTA 1174
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 70 GluProThrGlyCysValLeuTyrArgTyrGlySerPheSerThrGluLeuAsnIleVal 89
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1175 GAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGGTGCCATGG 1234
:|:| ||| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 90 GlnGlyIleGluAsnValAlaIleValGln-----ValValProAlaAlaProGlu 106
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1235 CCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGCCAAGGGAGCATTCACACGGAGGTC 1294
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 107 GlySerGlyAsnSerValGluLeuThrValThrCysGluGlySerLeuProGluGluVal 126
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1295 TGTACCATCATTTCTGACCCCCACCTGCGAGATCACCCAGAACACAGTCTGCGAGCCCTGTG 1354
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 127 CysThrValValAlaAlaAspAlaGluCysArgThrAlaGlnMetGlnThrCysSerAlaVal 146
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1355 GATGTGGATGAGATGTCTGTGACTGTGAGACGAAACCTTCAATGGGTCTGGGACGTAC 1414
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 147 AlaProAlaProGlyCysGlnLeuValLeuArgGlnAspPheAsnGlnSerGlyLeuTyr 166
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1415 TGTGTGAACCTCACCCCTGGGGATGACACAGCCTGGCTCTCACGAGACCCCTGATTTCT 1474
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 167 CysLeuAsnValSerLeuAlaAsnGlyAsnGlyLeuAlaValAlaSerThrArgValAla 186
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1475 GTTCCTGACAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCCTGATCTCCGTT 1534
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 187 ValGlyGlyAlaSerProAla-----AlaGlyGlyThrThrLeuThrVal 201
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1535 GGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAACACAGGAA 1594
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 202 GlyLeuLeuLeuIleValAlaAlaLeuGlyThrAlaAlaTyrThrTyrArgArgValLys 221
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1595 TACAAACCAATAGAAAAATAGTCCTGGGAATGTGTCAGA----- 1633
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 222 TyrSerProLeuLeuProThrAlaProGlnValSerArgProHisSerTrpLeuProPro 241
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1634 AGCAAAGGCCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTCCGGGA----- 1687
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 242 GlyAlaAlaLeuArgLeuLeuLeuArgGlnAla-----PheGlyGlyAlaPro 257
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1688 AACCAGGAAGGATCCGCTACTCAA 1714
:|:| ||| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 258 SerGlyGluSerSerProLeuLeuArg 266
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|

RESULT 13

Q9QY67

ID Q9QY67

AC Q9QY67;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

PRELIMINARY; PRT; 236 AA.

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mutant matrix melanosomal protein silver (Fragment).
GN SI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20063689; PubMed=10594241;
RA Martinez-Esparza M., Jimenez-Cervantes C., Bennet D.C., Lozano J.A.,
RA Solano F., Garcia-Borron J.C.;
RT "The mouse silver locus encodes a single transcript truncated by the
RT silver mutation.";
RL Mamm. Genome 10:1168-1171(1999).
DR EMBL; AF119092; AAF19761.1; -.
DR MGD; MGI:98301; SI.
DR InterPro; IPR007110; Ig-like.
FT NON TER 1
SQ SEQUENCE 236 AA; 24293 MW; 3ECB812D4DF10EE7 CRC64;

Alignment Scores:

Pred. No.: 7.98e-15 Length: 236
Score: 247.50 Matches: 62
Percent Similarity: 47.69% Conservative: 31
Best Local Similarity: 31.79% Mismatches: 85
Query Match: 5.10% Indels: 17
DB: 11 Gaps: 5

US-10-039-272-1 (1-2661) x Q9QY67 (1-236)

QY 1019 GGACCTTGTCCGCCACCGCCACCACCCAGACCTTCAAAACCCACCCCTTCTTTAGGA 1078
|||:|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 39 GlyProAspAlaSerProLeuLeuProThrGlnSerSer-----ThrGlySerIleSer 56
|||:|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1079 CCTGTGTGTGACAAACCCC-----CTGGAGCTGAGTAGGATTCCTGATGAAACCTGC 1129
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 57 ProLeuLeuAspAspThrAspThrIleMetLeuValLysArgGlnValProLeuAspCys 76
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1130 CAGATTAAACAGATATGGCCACTTTCAGAGCCACCATCACAAATTGTAGAGGAATCTTAGAG 1189
:|:| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 77 ValLeuTyrArgTyrGlySerPheSerLeuAlaLeuAspIleValGlnGly----- 93
||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1190 GTTAACATCATCCAGATGACAGACGTCCTGATGCCGTGCCATGGCCTGAAAGCTCCCTA 1249
|||:|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 94 -----IleGluSerAlaGluIleLeuGlnAlaValProSerSerGluGlyAspAla 110
|||:|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1250 ATAGACTTTTGTGTCGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACCATCATTTCT 1309
:|:| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 111 PheGluLeuThrValSerCysGlnGlyLeuProLysGluAlaCysMetAspIleSer 130
|||:|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1310 GACCCACCTCGGAGATCACCCAGAACACACAGTCTGCAGCCCTGTGGATGTGGATGAGATG 1369
|||:|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 131 SerProGlyCysGlnProProAlaGlnArgLeuCysGlnSerValProProSerProAsp 150
|||:|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1370 TGTCTGCTGACTGTGAGACGAAACCTTC---AATGGGTCTGGACGCTACTGTGTGAACCTC 1426
|||:|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 151 CysGlnLeuValLeuHisGlnValLeuLysGlySerGlyThrTyrCysLeuAsnVal 170
|||:|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1427 ACCCTGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCTCCTGACAGA 1486
:|:| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 171 SerLeuAlaAspAlaAsnSerLeuAlaValAlaSerThrGlnLeuValValProGlyGln 190
|||:|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1487 GACCCAGCCTCGCCTTTAAGGATGGCAACAGTCCCTGATCTCCGTTGGCTGGCC 1546
|||:|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 191 AspGlyGly-----LeuGlyGlnAlaProLeuLeuValGlyIleLeuLeu 205
|||:|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
QY 1547 ATATTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACAAAG 1591
:|:| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|
Db 206 ValLeuValAlaValValLeuAlaSerLeuIleHisArgHisArg 220
|||:|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:| ||| :|:|

RESULT 14

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 10, 2004, 15:32:29 ; Search time 41.5 Seconds
(without alignments)
6620.564 Million cell updates/sec

Title: US-10-039-272-1
Perfect score: 4850
Sequence: 1 cggcacgagggccagagga.....aaaaaaaaaaaaaaaaaaaa 2661

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10039272/runat_07092004_144224_21675/app_query.fasta_1.2823
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10039272 @CGN 1 1 48 @runat_07092004_144224_21675 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2999	61.8	560	1	US-09-985-799-90 Sequence 90, Appl
2	2999	61.8	560	1	US-08-594-031-90 Sequence 90, Appl
3	2999	61.8	560	4	US-09-643-597-225 Sequence 225, Appl
4	2999	61.8	560	4	US-09-480-884A-225 Sequence 225, Appl
5	2999	61.8	560	4	US-09-542-615A-225 Sequence 225, Appl
6	2999	61.8	560	4	US-09-606-421B-225 Sequence 225, Appl
7	2170	44.7	574	3	US-09-383-586-36 Sequence 36, Appl
8	2152	44.4	572	4	US-09-197-970B-5 Sequence 5, Appl
9	950.5	19.6	376	1	US-09-985-799-100 Sequence 100, Appl
10	950.5	19.6	376	1	US-09-985-799-102 Sequence 102, Appl
11	950.5	19.6	376	1	US-08-594-031-100 Sequence 100, Appl
12	950.5	19.6	376	1	US-08-594-031-102 Sequence 102, Appl

13	593	12.2	661	2	US-08-417-174-121	Sequence 121, Appl
14	593	12.2	661	3	US-09-267-439-121	Sequence 121, Appl
15	593	12.2	661	4	US-08-388-852B-2	Sequence 2, Appli
16	593	12.2	661	4	US-09-073-138-121	Sequence 121, Appl
17	592	12.2	661	2	US-08-417-174-27	Sequence 27, Appl
18	592	12.2	661	2	US-08-231-565A-27	Sequence 27, Appl
19	592	12.2	661	2	US-09-007-961-27	Sequence 27, Appl
20	592	12.2	661	3	US-09-267-439-27	Sequence 27, Appl
21	592	12.2	661	4	US-09-073-138-27	Sequence 27, Appl
22	574.5	11.8	668	1	US-07-891-942G-6	Sequence 6, Appli
23	238.5	4.9	190	4	US-08-388-852B-35	Sequence 35, Appl
24	236	4.9	202	4	US-08-388-852B-38	Sequence 38, Appl
25	233.5	4.8	192	4	US-08-388-852B-37	Sequence 37, Appl
26	233	4.8	197	4	US-08-388-852B-36	Sequence 36, Appl
27	124	2.6	71	1	US-09-985-799-98	Sequence 98, Appl
28	124	2.6	71	1	US-08-594-031-98	Sequence 98, Appl
29	111	2.3	525	4	US-09-976-594-64	Sequence 64, Appl
30	107.5	2.2	429	4	US-09-252-991A-31611	Sequence 31611, A
31	107	2.2	459	4	US-09-538-414-2	Sequence 2, Appli
32	107	2.2	459	4	US-10-074-279-2	Sequence 2, Appli
33	107	2.2	610	1	US-07-821-717B-6	Sequence 6, Appli
34	107	2.2	610	1	US-08-119-262B-6	Sequence 6, Appli
35	107	2.2	610	1	US-08-135-929A-11	Sequence 11, Appl
36	107	2.2	610	1	US-08-234-265A-11	Sequence 11, Appl
37	106.5	2.2	427	3	US-08-705-771-16	Sequence 16, Appl
38	106.5	2.2	427	4	US-09-417-540-16	Sequence 16, Appl
39	106.5	2.2	742	4	US-09-489-039A-7998	Sequence 7998, Ap
40	106	2.2	862	1	US-08-325-267A-4	Sequence 4, Appli
41	106	2.2	894	3	US-08-362-525-22	Sequence 22, Appl
42	106	2.2	894	3	US-08-971-692-15	Sequence 15, Appl
43	106	2.2	1537	1	US-08-325-267A-2	Sequence 2, Appli
44	104	2.1	865	4	US-09-281-766-19	Sequence 19, Appl
45	104	2.1	4302	3	US-08-658-136-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1

US-09-985-799-90
; Sequence 90, Application US/09985799
; Patent No. RE38392

GENERAL INFORMATION:

APPLICANT: THOMPSON, Timothy C.
TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
NUMBER OF SEQUENCES: 175
CORRESPONDENCE ADDRESS:

ADDRESSEE: BAKER & BOTTS, L.L.P.
STREET: 1299 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004-2400

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/985,799
FILING DATE: 06-No. RE38392-2001
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/594,031
FILING DATE: 30-JAN-1996
APPLICATION NUMBER: 60/006,838
FILING DATE: 16-NOV-1995

ATTORNEY/AGENT INFORMATION:

NAME: Remenick, James
REGISTRATION NUMBER: 36,902
REFERENCE/DOCKET NUMBER: 0A146-0110
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-639-7700
TELEFAX: 202-639-7890

```
;
;
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 560 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
;
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 90:
US-09-985-799-90

Alignment Scores:
Pred. No.:      1.86e-300      Length:      560
Score:          2999.00      Matches:      559
Percent Similarity: 99.82%      Conservative: 0
Best Local Similarity: 99.82%      Mismatches: 1
Query Match:      61.84%      Indels:      1
DB:              1      Gaps:      0

US-10-039-272-1 (1-2661) x US-09-985-799-90 (1-560)
QY 60 ATGGAATGCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTTCATGATGTGTGGCAATGAAAGACCTTCTCTGTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAATGGCTGGTCTTCTGATGATAAATGACTGGAATGAAAAAATCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAATACTCTGGAAGGAGGCCCGTGTGCGAGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80
QY 300 GTCCTGACCAAGTACTCACCAGCCCTCGTGGGCTCAAAATATAACATTTCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCTTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGCTGGTTTATCTCTGATCCATATGTTTACAACCTGGACAGCATGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGGAATAATGGCACCGGCCAAAGCCCATCAACGTCCTTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACCACCCGGATGGAGAAGATGGAATTTTCATCTACGTCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAAGTATTTCCAGAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAACAGATCAGATTCTCTGTGTTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
```

```
QY 780 ACTATGTTCCAGAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 250
QY 839 ATTATGTTTGATGTCCTGATTCATGATCCTAGCCACTTCTCCTCAATTATCTACCATTAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATACTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCCACCCAGACCTTCAAAACCCACCCCTTCTTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProProProProProSerLeuGly 340
QY 1079 CCTGCTGTGTGACAACCCCTGGAGCTGAGTAGGATTCTCTGATGAAAACTGCCAGATTAAAC 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTCAAGCCACCATCAACAATTGTAGAGGAATCTTAGAGGTTAACATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGCTCCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACCTGCCAAGGGAGCATTTCCCACGGAGGTCTGTACCATCATCTTGTGACCCCAAC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TCGGAGATCACCCAGAACACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACACAAGCCTGGCTCTCACGAGCACCCCTGATTCCTGTTCTGTTCTGTTCTGTTCTG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGGATGGCAAAACAGTGCCTGATCTCCGTTGGTGGTGGTGGTGGTGGTGGTGG 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTTGGTGTACAAAAACACAAAGGAATACAAACCAATAGAAAAATAGTCCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGGAAATGTGGTCAGAACCAAGGCCCTGAGTGTCTTCTCAACCCGTGCAAAAGCCGTGTC 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGGAACCCAGGAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAGGATTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 2
US-08-594-031-90
; Sequence 90, Application US/08594031
; Patent No. 5783182
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
```


; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,031
; FILING DATE: 30-JAN-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 560 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; US-08-594-031-90

Alignment Scores:
Pred. No.: 1.86e-300 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 1 Gaps: 0

US-10-039-272-1 (1-2661) x US-08-594-031-90 (1-560)

QY 60 ATGGAATGTTCTACTATTTCCTGGGATTTCTGCTCGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20

QY 120 GCCGCCAAACGATTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40

QY 180 CACAATCAATTAAATGGCTGTTCTTGATGAAAAATGACTGGAATGAAAAACTCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60

QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAATAACTCCTTGGAGGGAGGCCGTGTGAGGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80

QY 300 GTCCTGACCAGTCACTCACCAGCCCTCGTGGGCTCAAATATAACATTTCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100

QY 360 ATATTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120

QY 420 AGAAATGAGCGTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140

QY 480 GACAGTGACGGGAAAATGGCACCGGCCAAAGCCATCATCAACGTTCTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProAspGlyLys 160

QY 540 CCTTTTCTCACCAACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCTTCCACACACTT 599
Db 161 PropheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180

QY 600 GGTCAGTATTTCAGAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200

QY 660 GTGACACTTGGGCCTCAACTCATGGAAAGTAGCTGTCTACAGAAGACATGGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220

QY 720 GTTCCCATCGCACAAAGTCAAAGATGTGTACGTGTAACAGATCAGATTCCCTGTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240

QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCTATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerSerSerGluThrPheLeuLysAspLeuPro 260

QY 839 ATTATGTTTGATGTCCTGATTTCATGATCCTAGCCACTTCTCTCAATTATTCTTACCATTAA 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280

QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATACTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300

QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAAGTCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320

QY 1019 GGACCTTGTCCGCCACCGCCACCACCACCCAGACCTTCAAACCCACCCCTTCTTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProProProProProProProSer 340

QY 1079 CCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACCTGCCAGATTAA 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360

QY 1139 AGATATGGCCACTTTCAGCCACCACCATCACAAATTGTAGAGGGAATCTTAGAGTTAA 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380

QY 1199 ATCCAGATGACAGACGTCTCTGATGCCGTGCCATGGCCTGAAAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400

QY 1259 GTCGTGACCTGCCAAGGGAGCATTCACCGAGGAGTCTGTACCATCATTTCTGACCCACC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420

QY 1319 TGCAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440

QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGGACGCTACTGTGTGAACCTCACCCCTGGGG 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460

QY 1439 GACACAAGCCTGGCTCTCACGAGCACCCCTGATTTTCTGTTCTTGACAGAGACCCAGCCT 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480

QY 1499 CCTTTAAGGATGGCAAAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTGTCACT 1558

Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCTCTTGGTGTACAAAAACACAAGGAATACACCCCAATAGAAATAGTCCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGGAAATGCGTCAGAAAGCAAGGCCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTTT 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGGAACAGGAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 3

US-09-643-597-225
; Sequence 225, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-643-597-225

Alignment Scores:
Pred. No.: 1.86e-300 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 4 Gaps: 0

US-10-039-272-1 (1-2661) x US-09-643-597-225 (1-560)

QY 60 ATGGAATGTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTTCATGATGTCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTyrPheHisAspValLeuGlyAsnAspTyrPheHisAspValLeu 60
QY 240 GTGTGGAAGCGGGGAGACATGAGGTGAAAAAATCCTGGAAGGGAGGCCGTGTGCAGCG 299
Db 61 ValTyrLysArgGlyAspMetArgTyrLysAsnSerTyrLysGlyArgValGlnAla 80
QY 300 GTCCTGACCAAGTACTCACCAGCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100

QY 360 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGCTGGTTTATCTCTGATCCATATGTTTACAACTGGACAGCATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTACGGGGAAAATGGACCGGCCAAAGCCCATCATACCTCTTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProAspGlyLys 160
QY 540 CCTTTCTCTCACCACCCCGGATCGAGAAAGATGGAATTTTCATCTACGTCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTyrArgArgTyrAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAGTATTTCCAGAAAATGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCTCAACTCATGGAAGTGACTGTCTACAGAAAGATGGACGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTGATGTCCTGATTCATGATCCTAGCCACTTCTCCTCAATTTATCTACCATTAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCTGTTTGTGTTTCCACCAATCATACTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCGGCCACCGCCACCACCACCCAGACCTTCAAAACCCACCTTCTTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProProProProProProProPro 340
QY 1079 CCTGTGTTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACTGCCAGATTAAAC 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTCAAGCCACCACCATCACAAATGTAGAGGGAATCTTAGAGGTTAACATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGTCCTGATGCCGGTGCCATGCGCTGAAAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTyrProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACCATCATTTCTGACCCACC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TCGGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGTGATGTGTCTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460

QY 1439 GACACAAGCCTGGCTCTACAGACACCTGATTCTGTCTCCTGACAGAGACCAGCCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGGATGGCAACAGTCGCCCTGATCTCCGTTGGCTGCTTGGCCATATTGTCACT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTTGGTGTACAAAAACACACAGGAATACACCCCAATAGAAAAATAGTCCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGGAATGTGTCAGAAAGCAAGCCCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTTT 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGGAACAGGAAAGGATCCGCTACTCAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 4

US-09-480-884A-225
; Sequence 225, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-480-884A-225

Alignment Scores:
Pred. No.: 1.86e-300 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 4 Gaps: 0

US-10-039-272-1 (1-2661) x US-09-480-884A-225 (1-560)

QY 60 ATGGAATGCTCTACTATTCTCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAAAATCTTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGAAGCGGGAGACATGAGGTGGAAAAAATCCTTGGAGGGAGGCCGTGTGCAGGC 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80
QY 300 GTCCTGACCACTGACTCACAGCCCTCGTGGGCTCAAATATACATTTGCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100

QY 360 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGAAAAATGGCACCCGGCCCAAGCCCATCAACGCTTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACACACCCCGGATGGAGAAGATGGAATTTTCATCTACGCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCACTATTCCAGAAATTTGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCCTCAACTCATGGAAGTACTGTCTACAGAGACATGGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCTGTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTGATGTCCTGATTTCATGATCCTAGCCACTTCCCTCAATTTACTACCATTAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTGTGTTTCCACCAATCATACTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCCTTAACCTCACTGTGAAAGCTGACGACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCCGCCACCCACCCAGACCTTCAAAACCCACCCCTCTTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProProProProProProProPro 340
QY 1079 CCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTAAAC 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTCAAGCCACCATCACAAATGTAGAGGGAATCTTAGAGGTTAATATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGTCCTGATGCCGCTGCCATGGCCTGAAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACCTGCCAAGGGAGCATTCACCGGAGGTCTGTACCATCTTCTGACCCACC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TGCAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTCTCTGACAGAGACCCGCTCG 1498

Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGGATGGCAACAGTCCCTGATCTCCGTTGGTGGTGGCCATATTGTGCACT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCTCTTGGTGTAACAAAAACACAAGGAATACAAACCAATAGAAAATAGTCCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGAATGTGTCAGAAAGCCCTGAGTGTCTTTCTCAACCCGTGCAAAAGCCGTGTTTC 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGGAACACGAGAAAGGATCCCGTACTCAAAACCAAGATTAAAGGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 5
US-09-542-615A-225
; Sequence 225, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-542-615A-225

Alignment Scores:
Pred. No.: 1.86e-300 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 4 Gaps: 0

US-10-039-272-1 (1-2661) x US-09-542-615A-225 (1-560)

QY 60 ATGGAATGTCCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAAATGGCTGGTCTTCTGATGAAAATGACTGGAATGAAAACTCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTyrAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAAAACTCCTGGAAGGGAGGCGGTGTGCAGGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTyrLysAsnSerTrpLysGlyGlyArgValGlnAla 80
QY 300 GTCCTGACCAGTGTACACGACCCCTCGTGGCTCAATATAACATTTTGGCTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100

QY 360 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAATGAGGCTGGTTTATCTGCTGATCCATATGTTTAACTGGACAGCATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACACTGACGGGAAAAATGGACCGGCCAAAGCCATCATACCTCTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACCACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTyrArgArgTyrAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAGTATTTCCAGAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCCTCAACTCATGGAAAGTGAAGTGTCTACAGAAGACATGGACGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGACAAGTGAAGATGTGTACGTGGTGAACAGATCAGATTCTGTGTGTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAGAACGATCGAAATTTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTGATGCTCCTGATTTCATGATCCTAGCACCTTCTCAATTATTCTACCAATTAA 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTGTTTCCACCAATCATACTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTGCTCAATGGAAACCTTCAGCCCTTAACCTCACTGTGAAAGCTGCACACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCAACCCAGACCTTCAAAACCCACCCCTCTTTAGGA 1078
Db 321 GlyProCysProProProProProProProProArgProSerLysProThrProSerLeuGly 340
QY 1079 CCTGCTGGTGACAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTAAAC 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTCAAGCCCAACCATCACAAATTGTAGAGGAATCTTAGAGTTAAATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGTCCTGATGCGCGTGCATGGCCATGGCTGAAAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTyrProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACCTGCCAAGGGAGCATTCACCGGAGGTCTGTACCATCATTTCTGACCCCAACC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGTGGATGTGTCTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGACGCTACTGTGTGAACCTCACCTTGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACACAAGCCTTGCTCTCACGAGCACCCCTGATTTCTGTTCTCTGACAGACCCAGCCTCG 1498

Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGCTGGCCATATTGTCACT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTTGGTGTACAAAAACACAAGGAATACACCCCAATAGAAAATAGTCCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGAATGTGGTCAGAAAGCAAGGCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTC 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCGGGAAACAGGAAAGGATCCGCTACTCAAAAAACCAAGAAATTTAAAGGAGTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 6
US-09-606-421B-225
; Sequence 225, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRN
; ORGANISM: Homo sapien
US-09-606-421B-225

Alignment Scores:
Pred. No.: 1.86e-300 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 4 Gaps: 0

US-10-039-272-1 (1-2661) x US-09-606-421B-225 (1-560)

QY 60 ATGGAATGTCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTTCATGATGTCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAATAACTCCTGGAAGGAGCGGTGTGCAGGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80
QY 300 GTCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTG 359

Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTTACAACCTGGACAGCATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGAAAATGGCACCGGCCAAAAGCCATCAATAACGTCTTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACCAACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAGTATTTCCAGAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCCTCAACTCATGGAAAGTGACTGTCTACAGAAGACATGGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAGTGAAGATGTGTACGTGTAACAGATCAGATCAGATTCCTGTGTTGTG 779
Db 221 ValProIleAlaGlnVallLysAspValTyrValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAACGATCGAAATTTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTGATGTCCTGATTTCATGATCCTAGCCACTTTCCTCAATTATTCTACCATTAAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTAGCGCTGTTGTTTCCACCAATCATACTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGGAAACCTTTCAGCTTAACCTCAGTGTGAAAGCTGCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrVallLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCACCAACCCAGACACTTCAAAACCCACCCTTCTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProProProProProProProPro 340
QY 1079 CCTGCTGTGACAAACCCCTGGAGTGTAGTAGGATTTCCTGATGAAAACTGCCAGATTAAAC 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTCAAGCCACCACCATCACAAATTGTAGAGGGAATCTTAGAGGTTAACATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGTCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACCTGCCAAGGGAGCATTCACCGAGGTCGTGTACCATCATTTCTGACCCACC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGAT 1438

Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACACAAGCCTGGCTCTCACGAGCACCTGATTTCTGTCTCTGACAGAGACCCAGCCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGGATGGCAACAGTGGCCTGATCTCCGTTGGCTGCTGGCCATATTGTGCACT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTTGGTGTACAAAACACAAGGAATACAAACCCCAATAGAAAATAGTCCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGAATGTGTCAGAAAGCAAGGCCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTT 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGGAAACACGAGAAAGGATCCGCTACTCAAAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 7

US-09-383-586-36
; Sequence 36, Application US/09383586
; Patent No. 6242419
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Nevin
; APPLICANT: Onrust, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Murison, Greg
; TITLE OF INVENTION: Compounds isolated from stromal cells
; TITLE OF INVENTION: and methods for their use
; FILE REFERENCE: 11000.1037c1
; CURRENT APPLICATION NUMBER: US/09/383,586
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 36
; LENGTH: 574
; TYPE: PRT
; ORGANISM: Mouse
US-09-383-586-36

Alignment Scores:
Pred. No.: 6.87e-215 Length: 574
Score: 2170.00 Matches: 402
Percent Similarity: 82.63% Conservative: 69
Best Local Similarity: 70.53% Mismatches: 83
Query Match: 44.74% Indels: 17
DB: 3 Gaps: 2

US-10-039-272-1 (1-2661) x US-09-383-586-36 (1-574)

QY 60 ATGGAATGTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluSerLeuCysGlyValLeuGlyPheLeuLeuAlaAlaGlyLeuProLeuGln 20
QY 120 GCCGCCAAACGATTTCATGATGTCTGGCAATGAAAGACCTTCTGCTTACATGAGGAG 179
Db 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAsnHisMetArgGlu 40
QY 180 CACAATCAATTAAATGGCTGCTTCTGATGAAAATGACTGGAATGAAAACCTCTACCCA 239
Db 41 HisAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGluHisLeuTyrPro 60
QY 240 GTGTGAACGGGGAGACATGAGTGGAAAAAACTCCTGGAAGGAGGCCGTGTGCAGGCG 299
Db 61 ValTrpArgArgGlyAspGlyArgTrpLysAspSerTrpGluGlyGlyArgValGlnAla 80
QY 300 GTCCTGACCAGTGACTACACGACCTCGTGGGCTCAAATATAACATTTCGGGTGAACCTG 359

Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
QY 360 ATATTCCTTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAAGACTGC 419
Db 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGACAGCATGGTCAGAG 479
Db 121 ArgAsnAspLeuGlyLeuThrSerAspLeuHisValTyrAsnTrpThrAlaGlyAlaAsp 140
QY 480 GACAGTACGGGGAAAAATGGCACCGGCCAAAGCCATCATATAACGCTTTCCTGATGGAAA 539
Db 141 AspGlyAspTrpGluAspGlyThrSerArgSerGlnHisLeuArgPheProAspArg 160
QY 540 CTTTTCCTCACCAACCCCGATGGAGAAGATGGAATTTTCATCTACGTTCTCCACACACTT 599
Db 161 ProPheProArgProHisGlyTrpLysLysTrpSerPheValTyrValPheHisThrLeu 180
QY 600 GGTCAGTATTTCCAGAAAATTTGGACGATGTTTCTGAGAGATTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerAlaArgValSerIleAsnThrValAsn 200
QY 660 GTGACACTTGGGCTCAACTCATGGAACTGACTGTCTACAGAAAGACATGGACGGCATAT 719
Db 201 LeuThrAlaGlyProGlnValMetGluValThrValPheArgArgTyrGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTGTG 779
Db 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetSerGlnLysAsnAspArgAsnLeuSerAspGluIlePheLeuArgAspLeuPro 260
QY 839 ATTATGTTTGTATGCTCCTGATTCATGATTCCTAGCCACTTCTCAATTTTCTACCATTAAC 898
Db 261 IleValPheAspValLeuIleHisAspProSerHisPheLeuAsnAspSerAlaIleSer 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTAGCCCTTAACTGCTGTTTGTTCACCAATCATCTGTAAT 958
Db 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCCTTAACTCACTGTGAAAGCTGCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheAsnLeuAsnLeuThrValGlnThrAlaValPro 320
QY 1019 GGACCTTGTCCG-----CCACCGCCACCA 1042
Db 321 GlyProCysProProProSerProSerThrProProProSerThrProProSerPro 340
QY 1043 CCACCCAGACCT-----TCAAAACCCACCCCTTCTTTAGGACCTGCTGCTGAC 1090
Db 341 ProProSerProLeuProThrLeuSerThrProSerProSerLeuMetProThrGlyTyr 360
QY 1091 AACCCCTGGAGCTGAGTAGGATTCTGTGATGAAAACCTGCCAGATTAACAGATATGGCCAC 1150
Db 361 LysSerMetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyr 380
QY 1151 TTTCAAGCCACCATCACAATTTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACA 1210
Db 381 PheArgAlaThrIleThrIleValGluGlyIleLeuGluValSerIleMetGlnIleAla 400
QY 1211 GACGTCCTGATCCCGTGCCATGGCCTGAAAGCTCCCTAAATAGACTTTTGTGCTGACTGC 1270
Db 401 AspValProMetProThrProGlnProAlaAsnSerLeuMetAspPheThrValThrCys 420
QY 1271 CAAGGAGCATTTCCACGAGGCTGTGTACCATCATTTCTGACCCACCTCGGAGATCACC 1330
Db 421 LysGlyAlaThrProMetGluAlaCysThrIleIleSerAspProThrCysGlnIleAla 440
QY 1331 CAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTCTGTCTGCTGAGTGTGAGACGA 1390

Db 441 GlnAsnArgValCysSerProValAlaValAspGlyLeuCysLeuLeuSerValArgArg 460
QY 1391 ACCTTCATGGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGATGACACAAAGCCTG 1450
Db 461 AlaPheAsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAlaSerLeu 480
QY 1451 GCTCTCAGCAGCACCTGATTCTGTCTCTGACAGAGACCCAGCCTCGCCTTTAAGGATG 1510
Db 481 AlaLeuThrSerThrLeuIleSerIleProGlyLysAspProAspSerProLeuArgAla 500
QY 1511 GCAACAGTGCCTGATCTCCGTTGGTCTTGGCCATATTTGTCACTGTGATCTCCCTC 1570
Db 501 ValAsnGlyValLeuIleSerIleGlyCysLeuAlaValLeuValThrMetValThrIle 520
QY 1571 TTGGTGTACAAAAACACAAGGAATACAAACCAATAGAAAATAGCTCTGGGAATGTGGTC 1630
Db 521 LeuLeuTyrLysLysHisLysAlaTyrLysProIleGlyAsnCysProArgAsnThrVal 540
QY 1631 AGAAGCAAAAGGCCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGGAAC 1690
Db 541 LysGlyLysGlyLeuSerValLeuLeuSerHisAlaLysAlaProPhePheArgGlyAsp 560
QY 1691 CAGGAAAGGATCCGCTACTCAAAAACCAA 1720
Db 561 GlnGluLysAspProLeuLeuGlnAspLys 570

RESULT 8

US-09-197-970B-5
; Sequence 5, Application US/09197970B
; Patent No. 6664385
; GENERAL INFORMATION:
; APPLICANT: Michele Sanicola-Nadel
; Joseph V. Bonventre
; Catherine A. Hession
; Takaharu Ichimura
; Henry Wei
; Richard L. Cate
; TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Biogen, Inc.
; STREET: 14 Cambridge Center
; CITY: Cambridge
; STATE: MA
; COUNTRY: USA
; ZIP: 02142
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/197,970B
; FILING DATE: 23-No. 6664385-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/018,228
; FILING DATE: 24-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Levine, Leslie M.
; REGISTRATION NUMBER: 35,245
; REFERENCE/DOCKET NUMBER: A010 PCT CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 679-2810
; TELEFAX: (617) 679-2838
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 572 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:

US-09-197-970B-5

Alignment Scores:
Pred. No.: 4,94e-213 Length: 572
Score: 2152.00 Matches: 395
Percent Similarity: 82.57% Conservative: 74
Best Local Similarity: 69.54% Mismatches: 85
Query Match: 44.37% Indels: 15
DB: 4 Gaps: 1

US-10-039-272-1 (1-2661) x US-09-197-970B-5 (1-572)

QY 60 ATGGAATGTCTCTACTATTTCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluSerLeuCysGlyValLeuValPheLeuLeuLeuAlaGlyLeuProLeuGln 20
QY 120 GCCGCCAAACGATTTCATGATGTCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheArgAspValLeuGlyHisGluGlnTyrProAspHisMetArgGlu 40
QY 180 CACAATCAATTAAATGGCTGTCTTCTGATGAAATGACTGGAATGAAAACTCTACCCA 239
Db 41 AsnAsnGlnLeuArgGlyTrpSerSerAspGluAsnGluTrpAspGluGlnLeuTyrPro 60
QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAAAAACTCTCGAAGGGAGGCCGTGTGCAGGCG 299
Db 61 ValTrpArgArgGlyGluGlyArgTrpLysAspSerTrpGluGlyArgValGlnAla 80
QY 300 GTCCTGACCAGTCACTCACCAGCCCTCGTGGGCTCAAAATATAACATTTGCCGTGAACCTG 359
Db 81 AlaLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheValValAsnLeu 100
QY 360 ATATTCCTTAGATGCCAAAAGGAAGATGCCAATGSCAACATAGTCTATGAGAGAACTGC 419
Db 101 ValPheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluArgAsnCys 120
QY 420 AGAAATGAGGCTGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG 479
Db 121 ArgSerAspLeuGluLeuAlaSerAspProTyrValTyrAsnTrpThrThrGlyAlaAsp 140
QY 480 GACAGTGACGGGGAATAATGGCACCGGCCAAAGCCATCATACGTCCTCCCTGATGGAAA 539
Db 141 AspGluAspTrpGluAspSerThrSerGlnGlyGlnHisLeuArgPheProAspGlyLys 160
QY 540 CCTTTTCTCACACCCCGGATGGAGAGATGGAATTTTCATCTACGTCCTCCACACACTT 599
Db 161 ProPheProArgProHisGlyArgLysLysTrpAsnPheValTyrValPheHisThrLeu 180
QY 600 GGTCAGTATTTCCAGAAAATTGGACGATGTCAGTGAGAGATTTCTGTGAAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerAlaArgValSerIleAsnThrValAsn 200
QY 660 GTGACACTTGGCCTCAACTCATGGAGTACTGTCTACAGAAGACATGGACGGGCATAT 719
Db 201 LeuThrValGlyProGlnValMetGluValIleValPheArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTGTG 779
Db 221 IleProIleSerLysValLysAspValTyrValIleThrAspGlnIleProIlePheVal 240
QY 780 ACTATGTTCCAGAAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetTyrGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuArgAspLeuPro 260
QY 839 ATTATGTTTGATGTCCTGATTCATGATCCTAGCCACTTCCCTCAATTATTCTACCATTAAC 898
Db 261 IlePhePheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerAlaIleSer 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATACTGTGAAT 958
Db 281 TyrLysTrpAsnPheGlyAspAsnThrGlyLeuPheValSerAsnAsnHisThrLeuAsn 300
QY 959 CACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCA 1018

```
Db      301 HisThrTyrValLeuAsnGlyThrPheAsnPheAsnLeuThrValGlnThrAlaValPro 320
Qy      1019 GGACCTTGTCGCCACCGCCACCACCCACCCAGACCT----- 1054
Db      321 GlyProCysProSerProThrProSerProSerSerSerThrSerProSerProAlaSer 340
Qy      1055 -----TCAAAACCCACCCCTTCTTTAGGACCTGCTGGTGACAAACCCC 1096
Db      341 SerProSerProThrLeuSerThrProSerProSerProSerLeuMetProThrGlyHisLysSer 360
Qy      1097 CTGGAGCTGAGTAGGATTCCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAA 1156
Db      361 MetGluLeuSerAspIleSerAsnGluAsnCysArgIleAsnArgTyrGlyTyrPheArg 380
Qy      1157 GCCACCATCACAAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTC 1216
Db      381 AlaThrIleThrIleValAspGlyIleLeuGluValAsnIleIleGlnValAlaAspVal 400
Qy      1217 CTGATGCCGCTGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGTCGACCTGCCAAGGG 1276
Db      401 ProIleProThrProGlnProAspAsnSerLeuMetAspPheIleValThrCysLysGly 420
Qy      1277 AGCATTTCCACGAGGCTGTATCCATCATTTCTGACCCACCTCCGAGATCACCAGAAC 1336
Db      421 AlaThrProThrGluAlaCysThrIleIleSerAspProThrCysGlnIleAlaGlnAsn 440
Qy      1337 ACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGAACCTTC 1396
Db      441 ArgValCysSerProValAlaValAspGluLeuCysLeuLeuSerValArgAlaPhe 460
Qy      1397 AATGGGTCTGGGACGTACTGTGTGAACCTACCTCGGGGGATGACACAGCCTGGCTCTC 1456
Db      461 AsnGlySerGlyThrTyrCysValAsnPheThrLeuGlyAspAspAlaSerLeuAlaLeu 480
Qy      1457 ACGAGCACCCCTGATTTCTGTTCTTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAC 1516
Db      481 ThrSerAlaLeuIleSerIleProGlyLysAspLeuGlySerProLeuArgThrValAsn 500
Qy      1517 AGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTGTCACTGTGATCTCCCTCTTGGTG 1576
Db      501 GlyValLeuIleSerIleGlyCysLeuAlaMetPheValThrMetValThrIleLeuLeu 520
Qy      1577 TACAAAAAACACAAGGAATACAAACCAATAGAAAATAGTCCTGGGAATGTGTCAGAAAGC 1636
Db      521 TyrLysLysHisLysThrTyrLysProIleGlyAsnCysThrArgAsnValValLysGly 540
Qy      1637 AAAGGCCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGGAACCCAGGAA 1696
Db      541 LysGlyLeuSerValPheLeuSerHisAlaLysAlaProPheSerArgGlyAspArgGlu 560
Qy      1697 AAGGATCCGCTACTCAAAACCAA 1720
Db      561 LysAspProLeuLeuGlnAspLys 568
```

```
RESULT 9
US-09-985-799-100
; Sequence 100, Application US/09985799
; Patent No. RE38392
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
```

```
SOFTWARE: FastSEQ Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/985,799
FILING DATE: 06-No. RE38392-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/594,031
FILING DATE: 30-JAN-1996
APPLICATION NUMBER: 60/006,838
FILING DATE: 16-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Remenick, James
REGISTRATION NUMBER: 36,902
REFERENCE/DOCKET NUMBER: 0A146-0110
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-639-7700
TELEFAX: 202-639-7890
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 100:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
SEQUENCE DESCRIPTION: SEQ ID NO: 100:
US-09-985-799-100
Alignment Scores:
Pred. No.: 4e-89 Length: 376
Score: 950.50 Matches: 299
Percent Similarity: 59.31% Conservative: 26
Best Local Similarity: 54.56% Mismatches: 46
Query Match: 19.60% Indels: 177
DB: 1 Gaps: 61
US-10-039-272-1 (1-2661) x US-09-985-799-100 (1-376)
Qy      99 GCTGCAAGATTGCCACTTGATGCGCCCAACCAATTCATGATGTGCTGGGCAATGAAAGA 158
Db      6 AlaAlaArg-----AspAlaAlaLysArg---HisAspVal---GlyAsnArg--- 19
Qy      159 CCTTCTGCTTACATGAGGGAGCAACAATCAATTAAATGGCTGGTCTTCTGATGAAATGAC 218
Db      20 ---SerAlaTyrMetArg---HisAsn-----AsnGlyTrpSerSerAsp---AsnAsp 34
Qy      219 TGAATGAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGGAAAAACCTCTGG 278
Db      35 TrpAsnLysTyr-----ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrp 51
Qy      279 AAGGAGGCGGTGTGACCGGCTCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAT 338
Db      52 LysGlyGlyArgVal-----AlaValThrSerAspSer-----AlaValGlySerAsn 67
Qy      339 ATAACATTTGCGGTGACCTGATATTCCTAGATGCCAAAGGAAGATGCCAATGGCAAC 398
Db      68 Thr-----AlaValAsn-----ArgCys-----LysAspAlaAsnGlyAsn 79
Qy      399 ATAGTCTATGAGAAGAACTGCAGAAATAGAGCTGGTTATCTGCTGATCCATATGTTTAC 458
Db      80 ValTyr-----LysAsnCysArgAsn---AlaGly---SerAlaAsp---TyrValTyr 94
Qy      459 AACTGGACAGCATGTCAGAGACAGTGCACGGGAAATGGCACCGGCCAAAGCCATCAT 518
Db      95 AsnTrpThrAlaTrpSer---AspSerAspGly---AsnGlyThrGly---SerHisHis 111
Qy      519 AACGCTCTCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAAGATGGAATTC 578
Db      112 AsnVal-----AspGlyLys-----HisHisGlyTrpArgArgTrpAsnTrp 125
```


QY	159	CCTTCTGCTTACATGAGGAGCACAATCAATTAATGGCTGCTTCTGTGATGAAATGAC	218	Db	267	erAsp	-----ValValThrCys-----GlySer-----	T	275
Db	20	---SerAlaTyrMetArg---HisAsn-----AsnGlyTyrSerSerAsp---AsnAsp	34	QY	1299	CCATCATTTCTGACCCCACTTCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATG	1358	Db	275
QY	219	TGGATGAAAAAATCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGAAAACTCCTCG	278	Db	275	hrValCysThrSerAspThrCys-----ThrAsnThrValCysSer---ValAspV	291	QY	1359
Db	35	TrpAsnLysTyr-----ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrp	51	Db	291	alAsp---MetCys-----ThrValArgArgThr---AsnGlySerGlyThrTyrCysV	307	QY	1419
QY	279	AAGGAGCGCGTGTGCAGGCGTCTGACCACTGACTCACAGCCCTCGTGGGCTCAAT	338	Db	291	alAsp---MetCys-----ThrValArgArgThr---AsnGlySerGlyThrTyrCysV	307	QY	1419
Db	52	LysGlyGlyArgVal-----AlaValThrSerAspSer-----AlaValGlySerAsn	67	QY	1419	TGAACCTCACCTGGGGATGACACAAAGCCTGGCTCTCACGAGCACCTGATTCTGTTTC	1478	Db	307
QY	339	ATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAGGAAGATGCCAATGGCAAC	398	QY	1479	CTGACAGAGACCCAGCCTCGCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCT	1538	QY	1479
Db	68	Thr-----AlaValAsn-----ArgCys-----LysAspAlaAsnGlyAsn	79	Db	321	--AspArgAspAlaSer-----ArgMetAlaAsnSerAla-----SerValGlyC	335	QY	1539
QY	399	ATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTAC	458	Db	335	ys-----AlaValThrValSerValTyrLysLysHisLys-----	346	QY	1599
Db	80	ValTyr-----LysAsnCysArgAsn---AlaGly---SerAlaAsp---TyrValTyr	94	QY	1599	ACCCAATAGAAAAATAGTCTCGGAATGTGGTGCAGAAAGCAAGGCTGAGTGTCTTCTCA	1658	Db	347
QY	459	AACTGGACAGCATGTCAGAGGACAGTGACGGGGAATAATGGCACCGGCCAAAGCCATCAT	518	QY	1659	ACCGTGCAAAAGCCGTGTTCTTCCCGGGAACCCAGGAAAGGATCCGCTACTCAAAACC	1718	QY	1659
Db	95	AsnTrpThrAlaTrpSer---AspSerAspGly---AsnGlyThrGly---SerHisHis	111	Db	361	snArgAlaLysAlaVal-----GlyAsnLysAspLysAsn-----	372	QY	1719
QY	519	AACGTCTTCCCTGATGGGAAACCTTTTCTCACCAACCCCGGATGGAGAAGATGGAATTTC	578	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
Db	112	AsnVal-----AspGlyLys-----HisHisGlyTrpArgArgTrpAsnTyr	125	Db	373	-----LysGlyValSer	376	QY	1719
QY	579	ATCTACGTCTTCCACACACTTGGTGCAGTATTTCCAGAAATTTCCAGAAATTTCCAGT	638	Db	373	-----LysGlyValSer	376	QY	1719
Db	126	ValHisThrGlyTyrLys-----GlyArgCysSerValArg	137	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
QY	639	GTTTCTGTGAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAGTACTGTCTAC	698	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
Db	138	ValSerValAsnThrAlaAsnValThr-----GlyMetValThrValTyr	152	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
QY	699	AGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGACGTGGTAACA	758	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
Db	153	ArgArgHisGlyArgAlaTyrVal-----AlaValLysAspValTyrValValThr	169	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
QY	759	GATCAGATTCTCTGTGTGTGACTATGTTCCAGAGAAGCATCGAAATTCATCCGACGAA	818	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
Db	170	Asp-----ValValThrMet-----LysAsnAspArgAsnSerSerAsp--T	183	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
QY	819	ACCTTCCCAAGATCTCCCATATGTTTGTGATGCTCTGATTCATGATCCATCCACTTCC	878	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
Db	183	hr-----LysAspMetAsp-----ValHisAsp---SerHis-----	192	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
QY	879	TCAATTATTCTACCAATTAACATAAGTGAGCTTCGGGGATAATACTGGCCTGTTTGT	938	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
Db	193	--AsnTyrSerThr---AsnTyrLysTrpSer---GlyAspAsnThrGly-----ValS	208	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
QY	939	CCACCAATCATACTGTGAATCACACAGTATGTCTCAATGGAACCTTCAGCCTTAACCTCA	998	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
Db	208	erThrAsnHisThrValAsnHisThrTyrVal---AsnGlyThrSerAsn-----T	224	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
QY	999	CTGTGAAAGCTGCAGCACCCAGGACCTTGTCCGCCACCGCCACCCAGACCTTCAA	1058	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
Db	224	hrValLysAlaAlaAla-----GlyCys-----	231	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
QY	1059	AACCCACCCCTTCTTTAGGACCTGCTGTTGACAAACCCCTCGAGCTGAGTAGGATTCCTG	1118	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
Db	232	-----ArgSerLysThrSerGlyAlaGlyAspAsn-----SerArg-----	243	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
QY	1119	ATGAAAACCTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCACATCACAAATTGTAGAGG	1178	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
Db	244	--AspAsnCys-----AsnArgTyrGlyHis-----AlaThrThrValGly-	256	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
QY	1179	GAATCTTAGAGGTTAAATCATCATCCAGATGACAGACGTCTGATCCGGTGCCTG	1238	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
Db	257	-----ValAsnMetThrAspValMetVal-----TrpSerS	267	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373
QY	1239	AAAGCTCCCTAATAGACTTTGTCTGACCTTGCCAAAGGAGGATTTCCACGGAGTCTGTA	1298	QY	1719	AAGAAATTTAAAGGAGTTTCT	1738	Db	373

; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
US-08-594-031-100

Alignment Scores:
Pred. No.: 4e-89
Score: 950.50
Percent Similarity: 59.31%
Best Local Similarity: 54.56%
Query Match: 19.60%
DB: 1

US-10-039-272-1 (1-2661) x US-08-594-031-100 (1-376)

QY 99 GCTGCAAGATTGCCACTTGATGCGCCCAACGATTTTCATGATGTGCTGGCGCAATGAAAGA 158
Db 6 AlaAlaArg-----AspAlaAlaLysArg---HisAspVal---GlyAsnArg--- 19
QY 159 CCTTCTGCTTACATGAGGGAGGACAAATCAATTAAATGGCTGGTCTTCTGATGAAAAATGAC 218
Db 20 ---SerAlaTyrMetArg---HisAsn-----AsnGlyTrpSerSerAsp---AsnAsp 34
QY 219 TGAATGAAAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGGAATACTCCTGG 278
Db 35 TrpAsnLysTyr-----ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrp 51
QY 279 AAGGGAGGCGGTGTGCAGGCGGTCTCTGACCAGTGACTCACCAGCCCTCGTGGGCTCAAAT 338
Db 52 LysGlyGlyArgVal-----AlaValThrSerAspSer-----AlaValGlySerAsn 67
QY 339 ATAACATTTCGGGTGAACCTGATATTCCTAGATGCCAAAAAGGAAGATGCCAATGGCAAC 398
Db 68 Thr-----AlaValAsn-----ArgCys-----LysAspAlaAsnGlyAsn 79
QY 399 ATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTAC 458
Db 80 ValTyr-----LysAsnCysArgAsn---AlaGly---SerAlaAsp---TyrValTyr 94
QY 459 AACTGGACAGCATGTCAGAGGACAGTACGCGGGAATAATGGCACCGGCCCAAGCCATCAT 518
Db 95 AsnTrpThrAlaTrpSer---AspSerAspGly---AsnGlyThrGly---SerHisHis 111
QY 519 AACGTCTTCCTGATGGGAAACCTTTTCCTCACCACCCCGGATGGAGAACATGGAATTTC 578
Db 112 AsnVal-----AspGlyLys-----HisHisGlyTrpArgArgTrpAsnTyr 125
QY 579 ATCTACGTCTTCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGATGTTTCAGTGAGA 638
Db 126 ValHisThrGlyTyrLys-----GlyArgCysSerValArg 137
QY 639 GTTTCTGTGAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAGTGAAGTGTCTAC 698
Db 138 ValSerValAsnThrAlaAsnValThr-----GlyMetValThrValTyr 152
QY 699 AGAAGACATGGACGGGCATATGTTCCCATCGCACAAGTGAAGATGTGTACGTGTAACA 758
Db 153 ArgArgHisGlyArgAlaTyrVal-----AlaValLysAspValTyrValValThr 169
QY 759 GATCAGATTCTGTTGTTGTGACTATGTTCCAGAGAAGACGATCGAAATTCATCCGACGAA 818
Db 170 Asp-----ValValThrMet-----LysAsnAspArgAsnSerSerAsp--T 183
QY 819 ACCTTCCCAAGATCTCCCCATTATGTTGATGTCCTGATTCATTCATCCTAGCCACTTCC 878
Db 183 hr-----LysAspMetAsp-----ValHisAsp---SerHis----- 192
QY 879 TCAATTATTCTACCATTAACTACAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGT 938

Db 193 --AsnTyrSerThr-----AsnTyrLysTrpSer---GlyAspAsnThrGly-----Vals 208
QY 939 CCACCAATCATACTACTGTGAATCACACAGTATGTGCTCAATGGAACCTTACGCTTAACCTCA 998
Db 208 erThrAsnHisThrValAsnHisThrTyrVal---AsnGlyThrSerAsn-----T 224
QY 999 CTGTGAAAGCTGCAGACCAGGACCTGTGTCCGCCACCACCCAGCCAGACCTTCAA 1058
Db 224 hrValLysAlaAlaAla-----GlyCys----- 231
QY 1059 AACCCACCCCTTCTTTAGGACCTGCTGGTGACAACCCCTGGAGCTGAGTAGGATTCCTG 1118
Db 232 ----ArgSerLysThrSerGlyAlaGlyAspAsn-----SerArg----- 243
QY 1119 ATGAAAACTGCCAGATTAAACAGATATGCCACTTTTCAAGCCACCATCACAATTGTAGAGG 1178
Db 244 --AspAsnCys-----AsnArgTyrGlyHis-----AlaThrThrValGly- 256
QY 1179 GAATCTTAGAGGTTAACAATCATCCAGATGACAGACGTCCTGATGCCGGTGCCATGGCCTG 1238
Db 257 -----ValAsnMetThrAspValMetVal-----TrpSerS 267
QY 1239 AAAGCTCCCTAATAGACTTTGCTGACCTGCCAAGGGAGCAATTCACGAGGCTCTGTA 1298
Db 267 erAsp-----ValValThrCys---GlySer-----T 275
QY 1299 CCATCATTTCTGACCCCACTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATG 1358
Db 275 hrValCysThrSerAspThrCys-----ThrAsnThrValCysSer---ValAspV 291
QY 1359 TGGATGAGATGTGCTGCTGACTGTGAGACGAACTTCAATGGGTCTGGGACGTACTGTG 1418
Db 291 alAsp---MetCys-----ThrValArgArgThr---AsnGlySerGlyThrTyrCysV 307
QY 1419 TGAACCTCACCCCTGGGGATGACACAAAGCCTTGGCTCTCAGGAGCACCTGATTTCTGTT 1478
Db 307 alAsn-----ThrGlyAspAspThrSerAlaThrSerThrSerVal----- 320
QY 1479 CTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTCCCTGATCTCCGTTGGCT 1538
Db 321 --AspArgAspAlaSer-----ArgMetAlaAsnSerAla-----SerValGlyC 335
QY 1539 GCTTGGCCATATTGTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAAAGGAATACA 1598
Db 335 ys-----AlaValThrValSerValTyrLysLysHisLys----- 346
QY 1599 ACCCAATAGAAAATAGTCTCTGGGAATGGTTCAGAAAGCAAGGCTGAGTGTCTTCTCA 1658
Db 347 -----TyrAsnAsnSerGlyAsnValValArgSerLysGly-----SerValA 361
QY 1659 ACCGTGCAAAAGCCGTGTTCTTCCCGGAAACACGAGGAAAGGATCCGCTACTCAAAACC 1718
Db 361 snArgAlaLysAlaVal-----GlyAsnLysAspLysAsn----- 372
QY 1719 AAGAAATTAAAGGAGTTTCT 1738
Db 373 -----LysGlyValSer 376

RESULT 12
US-08-594-031-102
; Sequence 102, Application US/08594031
; Patent No. 5783182
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA

```

; ZIP: 20004-2400
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,031
; FILING DATE: 30-JAN-1996
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
;
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 102:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 376 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
;
; US-08-594-031-102

```

Alignment Scores:			
Pred. No.:	4e-89	Length:	376
Score:	950.50	Matches:	299
Percent Similarity:	59.31%	Conservative:	26
Best Local Similarity:	54.56%	Mismatches:	46
Query Match:	19.60%	Indels:	177
DB:	1	Gaps:	61

US-10-039-272-1 (1-2661) x US-08-594-031-102 (1-376)

QY	99	GCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTTCATGATGTGCTGGGCAATGAAAGA	158
Db	6	AlaAlaArg-----AspAlaAlaLysArg---HisAspVal---GlyAsnArg---	19
QY	159	CCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGGTCTTCTGTGATGAAAAATGAC	218
Db	20	---SerAlaTyrMetArg---HisAsn-----AsnGlyTrpSerSerAsp---AsnAsp	34
QY	219	TGGAATGAAAAAAGCTCTACCCAGTGTGGAAAGCGGGGAGACATGAGGTGGAAAAAACTCCTGG	278
Db	35	TrpAsnLysTyr-----ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrp	51
QY	279	AAGGAGGCCGCGTGTGCAGGCGGTCCCTGACCAGTGACTCACCGCCCTCGTGGGCTCAAAAT	338
Db	52	LysGlyGlyArgVal-----AlaValThrSerAspSer-----AlaValGlySerAsn	67
QY	339	ATAACATTTCGGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATGCCAATGGCAAC	398
Db	68	Thr-----AlaValAsn-----ArgCys-----LysAspAlaAsnGlyAsn	79
QY	399	ATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTAC	458
Db	80	ValTyr-----LysAsnCysArgAsn---AlaGly---SerAlaAsp---TyrValTyr	94
QY	459	AACTGGACAGCATGCTCAGAGGACAGTGCACGGGGAAAAATGGCACCGGCCAAAGCCCATCAT	518
Db	95	AsnTrpThrAlaTrpSer---AspSerAspGly---AsnGlyThrGly---SerHisHis	111
QY	519	AACGTCTTCCCTGATGGGAAAACCTTTTTCCTCACCAACCCCGGATGGAGAAGATGGAATTC	578

Db 347 -----TyrAsnAsnSerGlyAsnValValArgSerLysGly-----SerValA 361

QY 1659 ACCGTGCAAAAGCCGTGTTCTTCCCGGAAACACAGGAAAAGGATCCGCTACTCAAAAACC 1718

Db 361 snArgAlaLysAlaVal-----GlyAsnLysAspLysAsn----- 372

QY 1719 AAGAAATTTAAAGAGTTTCT 1738

Db 373 -----LysGlyValSer 376

RESULT 13

US-08-417-174-121

; Sequence 121, Application US/08417174

; Patent No. 5844075

; GENERAL INFORMATION:

; APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,

; APPLICANT: STEVEN A.

; TITLE OF INVENTION: MELANOMA ANTIGENS AND

; TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC

; TITLE OF INVENTION: METHODS

; NUMBER OF SEQUENCES: 126

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN, L.L.P.

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/417,174

; FILING DATE: 05-APR-1995

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/08/231,565

; FILING DATE: 22-APR-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: CAROL M. GRUPPI

; REGISTRATION NUMBER: 37,341

; REFERENCE/DOCKET NUMBER: 2026-4124US1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; TELEX: 421792

; INFORMATION FOR SEQ ID NO: 121:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 661

; TYPE: amino acid

; STRANDEDNESS: Unknown

; TOPOLOGY: Unknown

; MOLECULE TYPE: Protein

US-08-417-174-121

Alignment Scores:

Pred. No.: 4.31e-52 Length: 661

Score: 593.00 Matches: 167

Percent Similarity: 39.45% Conservative: 93

Best Local Similarity: 25.34% Mismatches: 199

Query Match: 12.23% Indels: 201

DB: 2 Gaps: 17

US-10-039-272-1 (1-2661) x US-08-417-174-121 (1-661)

QY 66 TGTCTCTACTATTTCCTGGGATTTCTGCTCGCTGGCTGCAAGATTGCCACTTGATGCCGCC 125

Db 8 CysLeu-----LeuHisLeuAlaValIleGlyAlaLeuLeuAlaValGlyAlaThr 24

QY 126 AAA-----CGATTTTCATGTGCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179

Db 25 LysValProArgAsnGlnAspTrpLeuGlyValSerArg----- 37

QY 180 CACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAATACTTACCCA 239

Db 38 -----GlnLeuArg-----ThrLysAlaTrpAsnArgGlnLeuTyrPro 50

QY 240 GTGTGGAAGCGGGGAGACATGAGGTGGAAAAAATCCTCTGGAAGGGAGCGGTGTGACGGC 299

Db 51 GluTrpThr-----GluAlaGlnArgLeuAspCysTrpArgGlyGlnValSerLeu 68

QY 300 GTCCTGACCAGTGACTCACCCAGCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTG 359

Db 69 LysValSerAsnAspGlyProThrLeuIleGlyAlaAsnAlaSerPheSerIleAlaLeu 88

QY 360 ATATTCCTTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419

Db 89 AsnPheProGlySerGlnLysValLeuProAspGlyGlnValIleTrpValAsnAsnThr 108

QY 420 AGAAATGAGGCTGTTTATCTGCTGATCCATATGTTTACAC-----TGG--- 464

Db 109 -----IleIleAsnGlySerGlnValTrpGly 117

QY 465 -----ACAGCATGGTCAGAGGACAGTGACGGGGAATAATGGCACCGGCCAAAGCCATCAT 518

Db 118 GlyGlnProValTyrProGlnGluThrAspAla----- 129

QY 519 AACGTCTTCCCTGATGGGAAACCTTTTCTCCTCACCCCGGATGGAGAAGATGGAAATTC 578

Db 130 CysIlePheProAspGlyGlyProCysProSerGlySerTrpSerGlnLysArgSerPhe 149

QY 579 ATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGATGTTTCAGTGAGA 638

Db 150 ValTyrValTrpLysThrTrpGlyGlnTyrTrpGlnValLeuGlyGlyProValSerGly 169

QY 639 GTTTCGTGTGAACACAGCCAATGTGACACTTGGGCCTCAACTCATGGAAGTGACTGTCTAC 698

Db 170 LeuSerIleGlyThrGlyArgAlaMetLeuGlyThrHisThrMetGluValThrValTyr 189

QY 699 AGAAGACATGGA---CGGGCATATGTTCCCATCGCACACAAGTGAAGATGTGTACGTGGTA 755

Db 190 HisArgArgGlySerArgSerTyrValProLeuAlaHisSerSerSerAlaPheThrIle 209

QY 756 ACAGATCAGATTCCTGTGTTGTGACTATGTCCAGAAAGAACGATCGAAATTCATCCGAC 815

Db 210 ThrAspGlnValProPheSerValSerValSerGlnLeuArgAlaLeuAspGlyGlyAsn 229

QY 816 GAAACCTTCCC-AAAGATCTCCCATTTATGTTTGTATGCTGCTGATTCATGATCCTAGCCAC 874

Db 230 LysHisPheLeuArgAsnGlnProLeuThrPheAlaLeuGlnLeuHisAspProSerGly 249

QY 875 TTCCTCAATTATTCTACCATTAACATAAGTGGAGCTTCGGGGATAATACTGGCCTGTTT 934

Db 250 TyrLeuAlaGluAlaAspLeuSerTyrThrTrpAspPheGlyAspSerSerGlyThrLeu 269

QY 935 GTTTCACCAATCATACTGTGAATCACACAGTATGTGCTCAATGGAACCTTCAGCCTTAAC 994

Db 270 IleSerArgAlaLeuValValThrHisThrTyrLeuGluProGlyProValThrAlaGln 289

QY 995 CTCACGTGAAAGCTGCAGCACCA-----GGACCTTGTCCGCCACCGCCA--- 1039

Db 290 ValValLeuGlnAlaAlaIleProLeuThrSerCysGlySerSerProValProGlyThr 309

QY 1039 ----- 1039

Db 310 ThrAspGlyHisArgProThrAlaGluAlaProAsnThrThrAlaGlyGlnValProThr 329

QY 1039 ----- 1039

Db 330 ThrGluValValGlyThrThrProGlyGlnAlaProThrAlaGluProSerGlyThrThr 349

QY 1039 ----- 1039

Db 350 SerValGlnValProThrThrGluValIleSerThrAlaProValGlnMetProThrAla 369

COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/267,439
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/417,174
FILING DATE: 05-APR-1995
APPLICATION NUMBER: US/08/231,565
FILING DATE: 22-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: CAROL M. GRUPPI
REGISTRATION NUMBER: 37,341
REFERENCE/DOCKET NUMBER: 2026-4124US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 758-4800
TELEFAX: (212) 751-6849
TELEX: 421792
INFORMATION FOR SEQ ID NO: 121:
SEQUENCE CHARACTERISTICS:
LENGTH: 661
TYPE: amino acid
STRANDEDNESS: Unknown
TOPOLOGY: Unknown
MOLECULE TYPE: Protein
US-09-267-439-121
Alignment Scores:
Pred. No.: 4,31e-52 Length: 661
Score: 593.00 Matches: 167
Percent Similarity: 39.45% Conservative: 93
Best Local Similarity: 25.34% Mismatches: 199
Query Match: 12.23% Indels: 201
DB: 3 Gaps: 17
US-10-039-272-1 (1-2661) x US-09-267-439-121 (1-661)
QY 66 TGTCTCTACTATTTCCTGGGATTTCCTGCTCCTGGCTGCAAGATTGCCACTTGTATGCGCGC 125
Db 8 CysLeu-----LeuHisLeuAlaValIleGlyAlaLeuLeuAlaValGlyAlaThr 24
QY 126 AAA-----CGATTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 25 LysValProArgAsnGlnAspTrpLeuGlyValserArg----- 37
QY 180 CACAATCAATTAATGGCTGGTCTTCTGTATGAAATGACTGGAATGAAAACTCTACCCA 239
Db 38 -----GlnLeuArg-----ThrLysAlaTrpAsnArgGlnLeuTrpPro 50
QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAAAAACTCTCGAAGGGAGGCGGTGTGCAGGCG 299
Db 51 GluTrpThr-----GluAlaGlnArgLeuAspCysTrpArgGlyGlnValSerLeu 68
QY 300 GTCCTGACCAGTGACTCACCGAGCCCTCGTGGGCTCAAAATATAAATGCGGTGAACCTG 359
Db 69 LysValSerAsnAspGlyProThrLeuIleGlyAlaAsnAlaSerPheSerIleAlaLeu 88
QY 360 ATATTCCTTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 89 AsnPheProGlySerGlnLysValLeuProAspGlyGlnValIleTrpValAsnAsnThr 108
QY 420 AGAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAAC-----TGG--- 464
Db 109 -----IleIleAsnGlySerGlnValTrpGly 117
QY 465 -----ACAGCATGGTCAGAGGACAGTACGGGAAAAATGGACCGGCCAAAGCCATCAT 518
Db 118 GlyGlnProValTrpProGlnGluThrAspAla----- 129
QY 519 AACGCTCTCCCTGATGGGAAACCTTTTCTCCTCACCCCGGATGGAGAGATGAATTTC 578
Db 130 CysIlePheProAspGlyGlyProCysProSerGlySerTrpSerGlnLysArgSerPhe 149

QY 1039 ----- 1039
Db 370 GluSerThrGlyMetThrProGluLysValProValSerGluValMetGlyThrThrLeu 389
QY 1039 ----- 1039
Db 390 AlaGluMetSerThrProGluAlaThrGlyMetThrProAlaGluValSerIleValVal 409
QY 1039 ----- 1039
Db 410 LeuSerGlyThrThrAlaAlaGlnValThrThrThrGluTrpValGluThrThrAlaArg 429
QY 1040 -----CCACCACCCAGACCTTCAAAACCC----- 1063
Db 430 GluLeuProIleProGluProGluGlyProAspAlaSerSerIleMetSerThrGluSer 449
QY 1064 ---ACCCCTCTTTAGGACCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGAT 1120
Db 450 IleThrGlySerLeuGlyProLeuLeuAspGlyThrAlaThrLeuArgLeuValLysArg 469
QY 1121 GAA-----AACTGCCAGATTAAACAGATATGGCCACTTTTCAAGCCACCATCACAAAT 1171
Db 470 GlnValProLeuAspCysValLeuTrpArgTyrGlySerPheSerValThrLeuAspIle 489
QY 1172 GTAGAGGGAATCTTAGAGGTTAATCATCATCATGACATGACAGACGCTCCTGATGCGGTGCCA 1231
Db 490 ValGlnGly-----IleGluSerAlaGluIleLeuGlnAlaValPro 503
QY 1232 TGGCCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGCCAAGGGAGCATTCACCGGAG 1291
Db 504 SerGlyGluGlyAspAlaPheGluLeuThrValSerCysGlnGlyGlyLeuProLysGlu 523
QY 1292 GTCTGTACCATCATTTCTGACCCACCTGCGAGATCACCCAGAACACAGTCTGCAGCCCT 1351
Db 524 AlaCysMetGluIleSerSerProGlyCysGlnProProAlaGlnArgLeuCysGlnPro 543
QY 1352 GTGGATGTGGATGATGTCTGTGCTGACTGTGAGACGAAACCTTC---AATGGGTCTGGG 1408
Db 544 ValLeuProSerProAlaCysGlnLeuValLeuHisGlnIleLeuLysGlyGlySerGly 563
QY 1409 ACGTACTGTGTGAACCTCACCCCTGGGGATGACACAAAGCTTGGCTCTCACGAGCACCCCTG 1468
Db 564 ThrTyrCysLeuAsnValSerLeuAlaAspThrAsnSerLeuAlaValValSerThrGln 583
QY 1469 ATTCTGTTCCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCCTGATC 1528
Db 584 LeuIleMetProGlyGlnGluAlaGly-----LeuGlyGlnValProLeuIle 599
QY 1529 TCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGACAAAAA 1585
Db 600 ValGlyIleLeuLeuValLeuMetAlaValValLeuAlaSerLeuIleTyrArgArg 618

RESULT 14
US-09-267-439-121
Sequence 121, Application US/09267439
Patent No. 6270778
GENERAL INFORMATION:
APPLICANT: KAWAKAMI, YUTAKA; ROSENBERG,
APPLICANT: STEVEN A.
TITLE OF INVENTION: MELANOMA ANTIGENS AND
TITLE OF INVENTION: THEIR USE IN DIAGNOSTIC AND THERAPEUTIC
TITLE OF INVENTION: METHODS
NUMBER OF SEQUENCES: 126
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN & FINNEGAN, L.L.P.
STREET: 345 PARK AVENUE
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10154
COMPUTER READABLE FORM:
MEDIUM TYPE: FLOPPY DISK

QY 579 ATCTACGTCITTCACACACACTTGGTCAGTATTTCCAGAAATTGGGACGATGTTTCAGTGAGA 638
Db 150 ValtYrValTrpLysThrTrpGlyGlnTyrTrpGlnValLeuGlyGlyProValSerGly 169
QY 639 GTTCTGTGAACACAGCCCAATGTGACACTTGGGCCTCAACTCATGGAAGTGAAGTGTCTAC 698
Db 170 LeuSerIleGlyThrGlyArgAlaMetLeuGlyThrHisThrMetGluValThrValTyr 189
QY 699 AGAAGACATGGA---CGGGCATATGTTCCCATCGCACAAGTGAAGATGTGACGTGGTA 755
Db 190 HisArgArgGlySerArgSerTyrValProLeuAlaHisSerSerSerAlaPheThrIle 209
QY 756 ACAGATCAGATTCTGTGTTTGTGACTATGTTCCAGAGAAGACGATCGAAATTCATCCGAC 815
Db 210 ThrAspGlnValProPheSerValSerValSerGlnLeuArgAlaLeuAspGlyGlyAsn 229
QY 816 GAAACCTTCCC-AAAGATCTCCCATTTATGTTGATGTCTCTGATTCATGATCCTAGCCAC 874
Db 230 LysHisPheLeuArgAsnGlnProLeuThrPheAlaLeuGlnLeuHisAspProSerGly 249
QY 875 TTCTCTCAATTATCTACCAATTAACATAAGTGGAGCTTCGGGGATAATACTGGCCTGTTT 934
Db 250 TyrLeuAlaGluAlaAspLeuSerTyrThrTrpAspPheGlyAspSerSerGlyThrLeu 269
QY 935 GTTTCACCAATCATACTGTAATCACACGTATGTGCTCAATGGAACCTTCAGCCTTAAC 994
Db 270 IleSerArgAlaLeuValValThrHisThrTyrLeuGluProGlyProValThrAlaGln 289
QY 995 CTCAGTGTGAAAGCTGCAGACCA-----GGACCTTGTCCGCCACCGCCA--- 1039
Db 290 ValValLeuGlnAlaAlaIleProLeuThrSerCysGlySerSerProValProGlyThr 309
QY 1039 ----- 1039
Db 310 ThrAspGlyHisArgProThrAlaGluAlaProAsnThrThrAlaGlyGlnValProThr 329
QY 1039 ----- 1039
Db 330 ThrGluValValGlyThrThrProGlyGlnAlaProThrAlaGluProSerGlyThrThr 349
QY 1039 ----- 1039
Db 350 SerValGlnValProThrThrGluValIleSerThrAlaProValGlnMetProThrAla 369
QY 1039 ----- 1039
Db 370 GluSerThrGlyMetThrProGluLysValProValSerGluValMetGlyThrThrLeu 389
QY 1039 ----- 1039
Db 390 AlaGluMetSerThrProGluAlaThrGlyMetThrProAlaGluValSerIleValVal 409
QY 1039 ----- 1039
Db 410 LeuSerGlyThrThrAlaAlaGlnValThrThrGluTrpValGluThrThrAlaArg 429
QY 1040 -----CCACCACCCAGACCTTCAAAACCC----- 1063
Db 430 GluLeuProIleProGluProGluProGluGlyProAspAlaSerSerIleMetSerThrGluSer 449
QY 1064 ---ACCCCTTCTTTAGGACCTGTGTGGTGACAACCCCTGGAGCTGAGTAGGATTCCTGAT 1120
Db 450 IleThrGlySerLeuGlyProLeuLeuAspGlyThrAlaThrLeuArgLeuValLysArg 469
QY 1121 GAA-----AACTGCCAGATTACAGATATGGCCACTTTCAGACCACCATCACAATT 1171
Db 470 GlnValProLeuAspCysValLeuTyrArgTyrGlySerPheSerValThrLeuAspIle 489
QY 1172 GTAGAGGGAATCTTAGAGGTTAATCATCATCCAGATGACAGACGCTCCTGATGCCGGTGCCA 1231
Db 490 ValGlnGly-----IleGluSerAlaGluIleLeuGlnAlaValPro 503

QY 1232 TGGCCTGAAAGCTCCCTAATAGACTTTTGTCTGACCTGCCAAGGGAGCATTCACGGAG 1291
Db 504 SerGlyGluGlyAspAlaPheGluLeuThrValSerCysGlnGlyGlyLeuProLysGlu 523
QY 1292 GTCTGTACCATCATTTCTGACCCCACTGCGAGATCACCCAGAACACACAGTCTGCAGCCCT 1351
Db 524 AlaCysMetGluIleSerSerProGlyCysGlnProProAlaGlnArgLeuCysGlnPro 543
QY 1352 GTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGAACCTTC---AATGGGTCTGG 1408
Db 544 ValLeuProSerProAlaCysGlnLeuValLeuHisGlnIleLeuLysGlyGlySerGly 563
QY 1409 ACGTACTGTGTGAACCTCACCCCTGGGGGATGACACAAAGCCTGCTCTCACGAGCACCCCTG 1468
Db 564 ThrTyrCysLeuAsnValSerLeuAlaAspThrAsnSerLeuAlaValValSerThrGln 583
QY 1469 ATTCTGTCTCTGACAGACAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTCCCTGATC 1528
Db 584 LeuIleMetProGlyGlnGluAlaGly-----LeuGlyGlnValProLeuIle 599
QY 1529 TCCGTTGGCTGTTGGCCATATTGTCACTGTGATCTCCCTCTTGGTGTACAAAAA 1585
Db 600 ValGlyIleLeuLeuValLeuMetAlaValValLeuAlaSerLeuIleTyrArgArg 618
RESULT 15
US-08-388-852B-2
; Sequence 2, Application US/08388852B
; Patent No. 6500919
; GENERAL INFORMATION:
; APPLICANT: Adema, Gosse Jan; Figdor, Carl Gustav.
; TITLE OF INVENTION: Melanoma associated antigenic polypeptide,
; TITLE OF INVENTION: epitopes thereof and vaccine against melanoma.
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Adema, Gosse Jan; Figdor, Carl Gustav
; STREET: Philips van Leydenlaan 25
; CITY: Nijmegen
; STATE: Brabant
; COUNTRY: the Netherlands
; ZIP: 6525 EX
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,852B
; FILING DATE: February 15, 1995
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 661 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-388-852B-2
Alignment Scores:
Pred. No.: 4.31e-52 Length: 661
Score: 593.00 Matches: 167
Percent Similarity: 39.45% Conservative: 93
Best Local Similarity: 25.34% Mismatches: 199
Query Match: 12.23% Indels: 201
DB: 4 Gaps: 17
US-10-039-272-1 (1-2661) x US-08-388-852B-2 (1-661)
QY 66 TGTCTCTACTATTTCCTGGGATTTCTGCTCGCTGCAAGATTGCCACTTGATGCCGCC 125
Db 8 CysLeu-----LeuHisLeuAlaValIleGlyAlaLeuLeuAlaValGlyAlaThr 24

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 10, 2004, 01:34:28 ; Search time 213 Seconds
(without alignments)
6932.978 Million cell updates/sec

Title: US-10-039-272-1
Perfect score: 2661
Sequence: 1 cggcacgagggccagagga.....aaaaaaaaaaaaaaaaa 2661

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2547.8	95.7	2669	1	US-09-985-799-91
2	2547.8	95.7	2669	1	US-09-985-799-101
3	2547.8	95.7	2669	1	US-08-594-031-91
4	2547.8	95.7	2669	1	US-08-594-031-101
5	1051.2	39.5	2303	4	US-09-197-970B-4
6	1047.8	39.4	2213	3	US-09-383-586-27
7	545.6	20.5	619	4	US-09-123-912-105
8	545.6	20.5	619	4	US-09-643-597-105
9	545.6	20.5	619	4	US-09-643-597-121
10	545.6	20.5	619	4	US-09-480-884A-105
11	545.6	20.5	619	4	US-09-480-884A-121
12	545.6	20.5	619	4	US-09-542-615A-105
13	545.6	20.5	619	4	US-09-542-615A-121
14	545.6	20.5	619	4	US-09-606-421B-105
15	545.6	20.5	619	4	US-09-606-421B-121
16	545.6	20.5	619	4	US-09-221-107-105
17	545.6	20.5	619	4	US-09-221-107-121
18	493	18.5	494	4	US-09-389-681-449
19	493	18.5	494	4	US-09-620-405B-449
20	493	18.5	494	4	US-09-433-826B-449
21	493	18.5	494	4	US-09-604-287A-449
22	493	18.5	494	4	US-09-834-759-449
23	480.8	18.1	698	3	US-09-040-984-5
24	480.8	18.1	698	4	US-09-123-912-5
25	480.8	18.1	698	4	US-09-643-597-5
26	480.8	18.1	698	4	US-09-480-884A-5
27	480.8	18.1	698	4	US-09-542-615A-5

28	480.8	18.1	698	4	US-09-606-421B-5	Sequence 5, Appli
29	480.8	18.1	698	4	US-09-221-107-5	Sequence 5, Appli
30	312.6	11.7	620	1	US-09-985-799-113	Sequence 113, App
31	312.6	11.7	620	1	US-08-594-031-113	Sequence 113, App
32	306.8	11.5	524	1	US-09-985-799-105	Sequence 105, App
33	306.8	11.5	524	1	US-08-594-031-105	Sequence 105, App
34	210.2	7.9	473	1	US-09-985-799-115	Sequence 115, App
35	210.2	7.9	473	1	US-08-594-031-115	Sequence 115, App
36	208.4	7.8	335	1	US-09-985-799-92	Sequence 92, Appl
37	208.4	7.8	335	1	US-09-985-799-94	Sequence 94, Appl
38	208.4	7.8	335	1	US-09-985-799-96	Sequence 96, Appl
39	208.4	7.8	335	1	US-09-985-799-123	Sequence 123, App
40	208.4	7.8	335	1	US-08-594-031-92	Sequence 92, Appl
41	208.4	7.8	335	1	US-08-594-031-94	Sequence 94, Appl
42	208.4	7.8	335	1	US-08-594-031-96	Sequence 96, Appl
43	208.4	7.8	335	1	US-08-594-031-123	Sequence 123, App
44	189.4	7.1	286	1	US-09-985-799-121	Sequence 121, App
45	189.4	7.1	286	1	US-08-594-031-121	Sequence 121, App

ALIGNMENTS

RESULT 1
US-09-985-799-91
; Sequence 91, Application US/09985799
; Patent No. RE38392
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/985,799
; FILING DATE: 06-No. RE38392-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,031
; FILING DATE: 30-JAN-1996
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: <Unknown>
; ORIGINAL SOURCE:
; SEQUENCE DESCRIPTION: SEQ ID NO: 91:
US-09-985-799-91

QY 867 CTAGCACTTCTCAATTAATCTACCAATTAATACTAAGTGGAGCTTCGGGATAATACTG 926
Db |||||
QY 900 CTAGCACTTCTCAATTAATCTACCAATTAATACTAAGTGGAGCTTCGGGATAATACTG 959
Db |||||
QY 927 GCCTGTTGTTTCCACCAATCATACTGTGAATCACACGTAATGTGCTCAATGGAACCTTCA 986
Db |||||
QY 960 GCCTGTTGTTTCCACCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1019
Db |||||
QY 987 GCCTTAACCTCACTGTGAAAGCTGCAGCACCAGGACCTTGTCCGACCCGACCCACCCAC 1046
Db |||||
QY 1020 GCCTTAACCTCACTGTGAAAGCTGCAGCACCAGGACCTTGTCCGACCCGACCCACCCAC 1079
Db |||||
QY 1047 CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGTGCTGGTGACACCCCTTGGAGCTGA 1106
Db |||||
QY 1080 CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGTGCTGGTGACACCCCTTGGAGCTGA 1139
Db |||||
QY 1107 GTAGGATTCCTGATGAAACCTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCA 1166
Db |||||
QY 1140 GTAGGATTCCTGATGAAACCTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCA 1199
Db |||||
QY 1167 CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGG 1226
Db |||||
QY 1200 CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGG 1259
Db |||||
QY 1227 TGCCATGGCCTGAAAGCTCCCTAATAGACTTTTGTGCTGACCTGCCAAGGAGCATTCCTCA 1286
Db |||||
QY 1260 TGCCATGGCCTGAAAGCTCCCTAATAGACTTTTGTGCTGACCTGCCAAGGAGCATTCCTCA 1319
Db |||||
QY 1287 CGGAGGTCTGTACCATCATTTCTGACCCCACTTCCGAGATCACCCAGAACACAGTCTGCA 1346
Db |||||
QY 1320 CGGAGGTCTGTACCATCATTTCTGACCCCACTTCCGAGATCACCCAGAACACAGTCTGCA 1379
Db |||||
QY 1347 GCCCTGTGGATGTGGATGAGATGTGTCTGTGCTGAGACGAAACCTTCAATGGGCTG 1406
Db |||||
QY 1380 GCCCTGTGGATGTGGATGAGATGTGTCTGTGCTGACTGTGAGACGAAACCTTCAATGGGCTG 1439
Db |||||
QY 1407 GGACGTACTGTGTGAACCTCACCCCTGGGGATGACAAAGCCTGGCTCTCACAGCACCC 1466
Db |||||
QY 1440 GGACGTACTGTGTGAACCTCACCCCTGGGGATGACAAAGCCTGGCTCTCACAGCACCC 1499
Db |||||
QY 1467 TGATTTCTGTTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCCTGA 1526
Db |||||
QY 1500 TGATTTCTGTTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCCTGA 1559
Db |||||
QY 1527 TCTCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGACAAAAAC 1586
Db |||||
QY 1560 TCTCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGACAAAAAC 1619
Db |||||
QY 1587 ACAAGGAATACAAACCAATAGAAATAGTCTGGGAATGTGTCAGAAAGCAAGGCCCTGA 1646
Db |||||
QY 1620 ACAAGGAATACAAACCAATAGAAATAGTCTGGGAATGTGTCAGAAAGCAAGGCCCTGA 1679
Db |||||
QY 1647 GTGCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCGGGAACCCAGGAAAGGATCCGC 1706
Db |||||
QY 1680 GTGCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCGGGAACCCAGGAAAGGATCCGC 1739
Db |||||
QY 1707 TACTCAAAACCAAGAAATTAAGGAGTTTCTTAAATTCGACCTGTTTCTGAAGCTCA 1766
Db |||||
QY 1740 TACTCAAAACCAAGAAATTAAGGAGTTTCTTAAATTCGACCTGTTTCTGAAGCTCA 1799
Db |||||
QY 1767 CTTTTAGTGCCATGTGATGTGAGATGTGTGAGTGGCTATTAAACCTTTTCTCTAAAG 1826
Db |||||
QY 1800 CTTTTAGTGCCATGTGATGTGAGATGTGTGAGTGGCTATTAAACCTTTTCTCTAAAG 1859
Db |||||
QY 1827 ATTATTGTTAAATAGATATTGTGGTTTGGGAAGTTGAATTTTTTATAGGTTAAATGTCA 1886
Db |||||
QY 1860 ATTATTGTTAAATAGATATTGTGGTTTGGGAAGTTGAATTTTTTATAGGTTAAATGTCA 1919
Db |||||
QY 1887 TTTTAGAGATGGGAGAGGGATTATCTGCAGCAGCTTCAGCCATGTTGTGAAACTGAT 1946
Db |||||
QY 1920 TTTTAGAGATGGGAGAGGGATTATCTGCAGCAGCTTCAGCCATGTTGTGAAACTGAT 1979
Db |||||

QY 1947 AAAAGCAACTAGCAAGGCTTCTTTTCATTATTTTATGTTTCACTTATAAAGTCTTAG 2006
Db |||||
QY 1980 AAAAGCAACTAGCAAGGCTTCTTTTCATTATTTTATGTTTCACTTATAAAGTCTTAG 2039
Db |||||
QY 2007 GTAACTAGTAGGATAGAAACACTGTGTCCGAGAGTAAGGAGAGAACTACTATTGATTA 2066
Db |||||
QY 2040 GTAACTAGTAGGATAGAAACACTGTGTCCGAGAGTAAGGAGAGAACTACTATTGATTA 2099
Db |||||
QY 2067 GAGCCTAAACCCAGGTTAACTGCAAGAAGAGGGGGATACCTTCAGCTTTCCATGTAACCTG 2126
Db |||||
QY 2100 GAGCCTAAACCCAGGTTAACTGCAAGAAGAGGGGGATACCTTCAGCTTTCCATGTAACCTG 2159
Db |||||
QY 2127 TATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAGCTAACTGAATCCCA 2186
Db |||||
QY 2160 TATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAGCTAACTGAATCCCA 2219
Db |||||
QY 2187 CTTCAATACACACTCATGAACCTCCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGA 2246
Db |||||
QY 2220 CTTCAATACACACTCATGAACCTCCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGA 2279
Db |||||
QY 2247 TGTGCACACTTGTCTAGACTCAGAAAAATACTACTCTCATAAATGGGTGGGAGTATTTTG 2306
Db |||||
QY 2280 TGTGCACACTTGTCTAGACTCAGAAAAATACTACTCTCATAAATGGGTGGGAGTATTTTG 2339
Db |||||
QY 2307 GTGACAACTTACTTGTCTGGCTGAGTGAAGGAATGATATTCATATATATTCATTATTCCCA 2366
Db |||||
QY 2340 GTGACAACTTACTTGTCTGGCTGAGTGAAGGAATGATATTCATATATATTCATTATTCCCA 2399
Db |||||
QY 2367 TGGACATTTAGTTAGTGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTGTGTA 2426
Db |||||
QY 2400 TGGACATTTAGTTAGTGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTGTGTA 2459
Db |||||
QY 2427 TATTTCCAAATTTTGTATAGTGTCTGCTGCACATATTTGAAATCAAAATATTAAGACTTTCC 2486
Db |||||
QY 2460 TATTTCCAAATTTTGTATAGTGTCTGCTGCACATATTTGAAATC-ATATATTAAGACTTTCC 2518
Db |||||
QY 2487 AAAAATTTGGTCCCTGGTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTG 2546
Db |||||
QY 2519 AAAGATGAGTCCCTGGTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTG 2578
Db |||||
QY 2547 GAACTAAACCAATTTACTATATGTTAGACAAGACATTTTTTTTTTCTCTGAAAAA 2606
Db |||||
QY 2579 TAACTAAACCAATCTACTATATGTTAGACATGACATTTCTCTCTCTCTGAAAAA 2638
Db |||||
QY 2607 -AAAATGAGGGAAGAGACAAAAA 2636
Db |||||
QY 2639 TAAAGTGTGGGAAGAGACAAAAA 2669
Db |||||

RESULT 3

US-08-594-031-91
; Sequence 91, Application US/08594031
; Patent No. 5783182
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/594,031
; FILING DATE: 30-JAN-1996
; CLASSIFICATION: 435

;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/006,838
; FILING DATE: 16-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Remenick, James
; REGISTRATION NUMBER: 36,902
; REFERENCE/DOCKET NUMBER: 0A146-0110
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
;
; INFORMATION FOR SEQ ID NO: 91:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; US-08-594-031-91

Query Match 95.7%; Score 2547.8; DB 1; Length 2669;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY	28	AACCTTGGTGCCTGCTGCCGTGAGAAATTCAGCATGGAATGCTCTACTATTTCCTGGGAT	87
Db	60	ACCTTGAGTGCCTGCGTCCGTGAGAAATTCAGCATGGAATGCTCTACTATTTCCTGGGAT	119
QY	88	TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCCAACGATTTTCATGATGCTGG	147
Db	120	TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCCAACGATTTTCATGATGCTGG	179
QY	148	GCAATGAAAGACCTTCTGCTTACATGAGGGAGCAATCAATTAATTAATGGCTGCTCTG	207
Db	180	GCAATGAAAGACCTTCTGCTTACATGAGGGAGCAATCAATTAATTAATGGCTGCTCTG	239
QY	208	ATGAAATGACTGGAATGAAAAAATCTACCCAGTGTGGAAGCGGGAGACATGAGGTGA	267
Db	240	ATGAAATGACTGGAATGAAAAAATCTACCCAGTGTGGAAGCGGGAGACATGAGGTGA	299
QY	268	AAACTCCTGGAAGGAGGCCGTGTGCAGGCGGTCTCTGACCACTGACTCACCAGCCCTCG	327
Db	300	AAACTCCTGGAAGGAGGCCGTGTGCAGGCGGTCTCTGACCACTGACTCACCAGCCCTCG	359
QY	328	TGGGCTCAAATATAACATTTGCGGTGAACCTGATATCCCTAGATGCCAAAGGAAGATG	387
Db	360	TGGGCTCAAATATAACATTTGCGGTGAACCTGATATCCCTAGATGCCAAAGGAAGATG	419
QY	388	CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC	447
Db	420	CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC	479
QY	448	CATATGTTTAACTGGACAGCATGGTCAGAGSACAGTGACGGGGAATAATGGCACCCGCC	507
Db	480	CATATGTTTAACTGGACAGCATGGTCAGAGSACAGTGACGGGGAATAATGGCACCCGCC	539
QY	508	AAAGCCATCAFAACGTCTTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAA	567
Db	540	AAAGCCATCAFAACGTCTTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAA	599
QY	568	GATGGAATTTCACTACGTCTTCCACACACATTTGGTCAGTATTTCCAGAAATTTGGACGAT	627
Db	600	GATGGAATTTCACTACGTCTTCCACACACATTTGGTCAGTATTTCCAGAAATTTGGACGAT	659
QY	628	GTTTCAGTGAGATTCTGTGAACACACAGCCAAATGTGACACTTTGGGCCCTCAACTCATGGAAG	687
Db	660	GTTTCAGTGAGATTCTGTGAACACACAGCCAAATGTGACACTTTGGGCCCTCAACTCATGGAAG	719
QY	688	TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAGTGAAGATGTGT	747

Db	720	TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAGTGAAGATGTGT	779
QY	748	ACGTGGTAAACAGATCAGATTCCCTGTGTTGTGACTATGTTCCAGAAGACGATCGAAATT	807
Db	780	ACGTGGTAAACAGATCAGATTCCCTGTGTTGTGACTATGTTCCAGAAGACGATCGAAATT	839
QY	808	CATCCGACGAAACCTTCC-CAAAGATCTCCCCATTATGTTTGATGCTCCTGATTCATGATC	866
Db	840	CATCCGACGAAACCTTCCCTCAAAGATCTCCCCATTATGTTTGATGCTCCTGATTCATGATC	899
QY	867	CTAGCCACTTCTCAATTATTCTACCAATAACTACAAGTGGAGCTTCGGGGATAACTG	926
Db	900	CTAGCCACTTCTCAATTATTCTACCAATAACTACAAGTGGAGCTTCGGGGATAACTG	959
QY	927	GCCTGTTTGTTCACCAATCATACTGTGAATCACACGATATGCTCAATGGAACCTTCA	986
Db	960	GCCTGTTTGTTCACCAATCATACTGTGAATCACACGATATGCTCAATGGAACCTTCA	1019
QY	987	GCCTTAAACCTCACTGTGAAAGCTGCAGCACACGAGGACCTTGTCCGCCACCGCCACCAC	1046
Db	1020	GCCTTAAACCTCACTGTGAAAGCTGCAGCACACGAGGACCTTGTCCGCCACCGCCACCAC	1079
QY	1047	CCAGACCTTCAAACCCCTTCTTTTAGGACCTGTGCTGTCACAAACCCCTCGAGCTGA	1106
Db	1080	CCAGACCTTCAAACCCCTTCTTTTAGGACCTGTGCTGTCACAAACCCCTCGAGCTGA	1139
QY	1107	GTAGGATTCTGATGAAATGTCAGATTAACAGATATGSCACATTTCAAGCCACCATCA	1166
Db	1140	GTAGGATTCTGATGAAATGTCAGATTAACAGATATGSCACATTTCAAGCCACCATCA	1199
QY	1167	CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGG	1226
Db	1200	CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGG	1259
QY	1227	TGCCATGGCCTGAAAGCTCCCTAATAGACTTTTGTGCTGACCTGCCAAGGGAGCATTTCCCA	1286
Db	1260	TGCCATGGCCTGAAAGCTCCCTAATAGACTTTTGTGCTGACCTGCCAAGGGAGCATTTCCCA	1319
QY	1287	CGGAGGTCTGTACCATCATTTCTGACCCCTGCGGAGATCACCCAGAACACAGTCTGCA	1346
Db	1320	CGGAGGTCTGTACCATCATTTCTGACCCCTGCGGAGATCACCCAGAACACAGTCTGCA	1379
QY	1347	GCCCTGTGGATGGATGAGATGTGTCTGTGACTGTGAGACGAACCTTCAATGGGTCTG	1406
Db	1380	GCCCTGTGGATGGATGAGATGTGTCTGTGACTGTGAGACGAACCTTCAATGGGTCTG	1439
QY	1407	GGACGTACTGTGAAACCTCACCCCTGGGGATGACACAAGCCTGGCTCTCAGGACACCC	1466
Db	1440	GGACGTACTGTGAAACCTCACCCCTGGGGATGACACAAGCCTGGCTCTCAGGACACCC	1499
QY	1467	TGATTTCTGTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCTGA	1526
Db	1500	TGATTTCTGTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCTGA	1559
QY	1527	TCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAC	1586
Db	1560	TCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAC	1619
QY	1587	ACAAGGAATACACCCCAATAGAAAATAGTCTTGGGAATGTGGTCAGAAACAAAGGCTGA	1646
Db	1620	ACAAGGAATACACCCCAATAGAAAATAGTCTTGGGAATGTGGTCAGAAACAAAGGCTGA	1679
QY	1647	GTGTCTTTCTCAACCGTGCAAAAGCGTGTCTTCTCCCGGAAACAGGAAAAGGATCCGC	1706
Db	1680	GTGTCTTTCTCAACCGTGCAAAAGCGTGTCTTCTCCCGGAAACAGGAAAAGGATCCGC	1739
QY	1707	TACTCAAAAAACAAGAAATTTAAAGAGTTTCTTAAATTTTCGACCTTGTCTTGAAGCTCA	1766
Db	1740	TACTCAAAAAACAAGAAATTTAAAGAGTTTCTTAAATTTTCGACCTTGTCTTGAAGCTCA	1799
QY	1767	CTTTTCAGTGCCATTGTATGTGAGATGTGCTGGAGTGGCTATTAACTTTTTCCTAAAG	1826

Db 1800 CTTTCAGTGCCATTGATGTGAGATGTGCTGGAGTGGCTATTAAACCTTTTTTCCCTAAAG 1859
QY 1827 ATTATTGTTAAATAGATATTGTGTTTGGGGAAGTTGAATTTTTTATAGTTTAAATGTCA 1886
Db 1860 ATTATTGTTAAATAGATATTGTGTTTGGGGAAGTTGAATTTTTTATAGTTTAAATGTCA 1919
QY 1887 TTTTAGAGTGGGAGAGGGATTATATCTGCAGGCAGCTTCAGCCACTGTTGTGAAACTGAT 1946
Db 1920 TTTTAGAGATGGGAGAGGGATTATATCTGCAGGCAGCTTCAGCCACTGTTGTGAAACTGAT 1979
QY 1947 AAAAGCAACTTAGCAAGGCTTCTTTTCATTTATTTTATGTTTCACTTATAAAGCTTTAG 2006
Db 1980 AAAAGCAACTTAGCAAGGCTTCTTTTCATTTATTTTATGTTTCACTTATAAAGCTTTAG 2039
QY 2007 GTAACCTAGTAGAGAGAAACACTGTGTCGAGAGAGTAAAGGAGAGAGTACTATTGATTA 2066
Db 2040 GTAACCTAGTAGAGAGAAACACTGTGTCGAGAGAGTAAAGGAGAGAGTACTATTGATTA 2099
QY 2067 GAGCCTAACCCAGGTTAACTGCAAGAGAGGGGGATACCTTTCAGCTTTCATGTAACCTG 2126
Db 2100 GAGCCTAACCCAGGTTAACTGCAAGAGAGGGGGATACCTTTCAGCTTTCATGTAACCTG 2159
QY 2127 TATGCATAAAGCCCAATGTAGTCCAGTTTCTTAAGATCATGTTTCCAAAGCTAACTGAATCCCA 2186
Db 2160 TATGCATAAAGCCCAATGTAGTCCAGTTTCTTAAGATCATGTTTCCAAAGCTAACTGAATCCCA 2219
QY 2187 CTTCAATACACACTCATGAACCTCCTGATGGAACAATAACAGGCCCAAGCCTGTGTTATGA 2246
Db 2220 CTTCAATACACACTCATGAACCTCCTGATGGAACAATAACAGGCCCAAGCCTGTGTTATGA 2279
QY 2247 TGTGCACACTTGTCTAGACTCAGAAAAAATACTACTCTCATAAATGGTGGGAGTATTTG 2306
Db 2280 TGTGCACACTTGTCTAGACTCAGAAAAAATACTACTCTCATAAATGGTGGGAGTATTTG 2339
QY 2307 GTGACAAACCTACTTTGCTGGCTGAGTGAAGGAATGATATTATATATTCAATTATTTCCA 2366
Db 2340 GTGACAAACCTACTTTGCTGGCTGAGTGAAGGAATGATATTATATATTCAATTATTTCCA 2399
QY 2367 TGGACATTTAGTTAGTCTTTTATATACCGGCATGATGCTGAGTGACACTCTTGTGTA 2426
Db 2400 TGGACATTTAGTTAGTCTTTTATATACCGGCATGATGCTGAGTGACACTCTTGTGTA 2459
QY 2427 TATTTCCAAATTTTGTATAGTCTGACATATTTTGAATCAAAATATTAAGACTTTCC 2486
Db 2460 TATTTCCAAATTTTGTATAGTCTGACATATTTTGAATC-ATATATTAAGACTTTCC 2518
QY 2487 AAAAAATTTGGTCCCTGTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTG 2546
Db 2519 AAAGATGAGGTCCCTGTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTG 2578
QY 2547 GAACTAAAAACCATTTACTATATGTAGACAAGACATTTTTTTTTTTTCTTCCCTGTAAGAA 2606
Db 2579 TAACTAAAAACCATCTACTATATGTAGACATGACATTTCTTTTCTCTCTCTCTCTCTCTCT 2638
QY 2607 -AAATGAGGGAAGAGACAAAAA 2636
Db 2639 TAAAGTGGGAAGAGACAAAAA 2669

RESULT 4
US-08-594-031-101
; Sequence 101, Application US/08594031
; Patent No. 5783182
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Timothy C.
; TITLE OF INVENTION: METHOD FOR IDENTIFYING METASTATIC SEQUENCES
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; City: Washington
; STATE: DC
; COUNTRY: USA

ZIP: 20004-2400
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/594,031
FILING DATE: 30-JAN-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/006,838
FILING DATE: 16-NOV-1995
ATTORNEY/AGENT INFORMATION:
NAME: Remenick, James
REGISTRATION NUMBER: 36,902
REFERENCE/DOCKET NUMBER: 0A146-0110
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-639-7700
TELEFAX: 202-639-7890
TELEX:
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 2669 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE:
ORIGINAL SOURCE:
US-08-594-031-101

Query Match 95.7%; Score 2547.8; DB 1; Length 2669;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;
QY 28 AACCTTGGTGCCTGCTCCGTGAGAATTTCAGCATGGAATGTCTCTACTATTTCCTGGGAT 87
Db 60 ACCTTGAGTGCCCTGCTCCGTGAGAATTTCAGCATGGAATGTCTCTACTATTTCCTGGGAT 119
QY 88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCGGCCCAACGATTTTCATGATGTGCTGG 147
Db 120 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCGGCCCAACGATTTTCATGATGTGCTGG 179
QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCAATCAATTAATGGCTGCTCTCTG 207
Db 180 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCAATCAATTAATGGCTGCTCTCTG 239
QY 208 ATGAAAAATGACTGGAATGAAAAAATCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 267
Db 240 ATGAAAAATGACTGGAATGAAAAAATCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 299
QY 268 AAAACTCCTGGAAGGGAGGCGGTGTGCGAGGCGGTCTGACCAAGTACTCACCAGCCCTCG 327
Db 300 AAAACTCCTGGAAGGGAGGCGGTGTGCGAGGCGGTCTGACCAAGTACTCACCAGCCCTCG 359
QY 328 TGGGCTCAAATATAACATTTGCGGTGAACCTGATATTTCCCTAGATGCCAAAAGGAAGATG 387
Db 360 TGGGCTCAAATATAACATTTGCGGTGAACCTGATATTTCCCTAGATGCCAAAAGGAAGATG 419
QY 388 CCAATGGCAACATAGTCTATGAAAGAACTGCAGAAATGAGCTGGTTTATCTGCTGATC 447
Db 420 CCAATGGCAACATAGTCTATGAAAGAACTGCAGAAATGAGCTGGTTTATCTGCTGATC 479
QY 448 CATATGTTTACAACACTGGACAGCATGGTTCAGAGGACAGTACGGGGAATAATGGCACCGGCC 507
Db 480 CATATGTTTACAACACTGGACAGCATGGTTCAGAGGACAGTACGGGGAATAATGGCACCGGCC 539
QY 508 AAAGCCCATATAACGTTCTTCCCTGATGGGAAACCTTTTCTCACACCCCGGATGGAGAA 567
Db 540 AAAGCCCATATAACGTTCTTCCCTGATGGGAAACCTTTTCTCACACCCCGGATGGAGAA 599

QY 568 GATGGAATTTCACTACGCTCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGAT 627
Db |||||
QY 600 GATGGAATTTCACTACGCTCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGAT 659
Db |||||
QY 628 GTTCAGTGAGAGTTTCTGTGAACACAGACCAATGTGACACTTGGGCTCAACTCATGGAAG 687
Db |||||
QY 660 GTTCAGTGAGAGTTTCTGTGAACACAGACCAATGTGACACTTGGGCTCAACTCATGGAAG 719
Db |||||
QY 688 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCAAAAGTGAAGATGTGT 747
Db |||||
QY 720 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCAAAAGTGAAGATGTGT 779
Db |||||
QY 748 ACGTGGTAAACAGATCAGATTCCTGTGTGTTGTGACTATGTTCCAGAAACGATCGAAAT 807
Db |||||
QY 780 ACGTGGTAAACAGATCAGATTCCTGTGTGTTGTGACTATGTTCCAGAAACGATCGAAAT 839
Db |||||
QY 808 CATCCGACGAAACCTTCC-CAAGATCTCCCCATTATGTTGATGTCCTGATTCATGATC 866
Db |||||
QY 840 CATCCGACGAAACCTTCCCTCAAGATCTCCCCATTATGTTGATGTCCTGATTCATGATC 899
Db |||||
QY 867 CTAGCCACTTCTCTCAATATTCTACCATTAACATAAGTGGAGCTTCGGGGATAATACTG 926
Db |||||
QY 900 CTAGCCACTTCTCTCAATATTCTACCATTAACATAAGTGGAGCTTCGGGGATAATACTG 959
Db |||||
QY 927 GCCTGTTGTTTCCACCATCATACTGTGAATCACACGATGTCCTCAATGGAACCTTCA 986
Db |||||
QY 960 GCCTGTTGTTTCCACCATCATACTGTGAATCACACGATGTCCTCAATGGAACCTTCA 1019
Db |||||
QY 987 GCCTTAACCTCACTGTGAAAGTGCAGCACCGACCTGTCCGCCACCGCCACCAAC 1046
Db |||||
QY 1020 GCCTTAACCTCACTGTGAAAGTGCAGCACCGACCTGTCCGCCACCGCCACCAAC 1079
Db |||||
QY 1047 CCAGACCTTCAAAACCCACCTTCTTTAGGACCTGTGGTGACAAACCCCTGAGCTGA 1106
Db |||||
QY 1080 CCAGACCTTCAAAACCCACCTTCTTTAGGACCTGTGGTGACAAACCCCTGAGCTGA 1139
Db |||||
QY 1107 GTAGGATTCTCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCA 1166
Db |||||
QY 1140 GTAGGATTCTCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCA 1199
Db |||||
QY 1167 CAATTGTAGAGGGAATCTTAGAGGTTAAACATATCCAGATGACAGACGTCCTGATGCCGG 1226
Db |||||
QY 1200 CAATTGTAGAGGGAATCTTAGAGGTTAAACATATCCAGATGACAGACGTCCTGATGCCGG 1259
Db |||||
QY 1227 TGCCATGGCCTGAAAGCTCCCTAAATAGACTTTTGTGCGTACCTGCCAAGGGAGCATTTCCA 1286
Db |||||
QY 1260 TGCCATGGCCTGAAAGCTCCCTAAATAGACTTTTGTGCGTACCTGCCAAGGGAGCATTTCCA 1319
Db |||||
QY 1287 CGGAGGTCTGTACCATCATTTCTGACCCCACTGCGAGATCAACCCAGAAACAGTCTGCA 1346
Db |||||
QY 1320 CGGAGGTCTGTACCATCATTTCTGACCCCACTGCGAGATCAACCCAGAAACAGTCTGCA 1379
Db |||||
QY 1347 GCCCTGTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGAAACCTTCAATGGGTCTG 1406
Db |||||
QY 1380 GCCCTGTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGAAACCTTCAATGGGTCTG 1439
Db |||||
QY 1407 GGACGTACTGTGTGAACCTCACCCCTGGGGATGACACAAGCCTGGCTCTCACGAGCACCC 1466
Db |||||
QY 1440 GGACGTACTGTGTGAACCTCACCCCTGGGGATGACACAAGCCTGGCTCTCACGAGCACCC 1499
Db |||||
QY 1467 TGATTTCTGCTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCTGA 1526
Db |||||
QY 1500 TGATTTCTGCTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCTGA 1559
Db |||||
QY 1527 TCTCCGTTGGCTGCTTGGCCATATTTGTCTACTGTGATCTCCCTCTTGTGTACAAAAAC 1586
Db |||||
QY 1560 TCTCCGTTGGCTGCTTGGCCATATTTGTCTACTGTGATCTCCCTCTTGTGTACAAAAAC 1619
Db |||||
QY 1587 ACAAGGAATACAAACCAATAGAAAAATAGTCTCTGGGAATGTGGTCAAGAACCAAGGCCTGA 1646
Db |||||
QY 1620 ACAAGGAATACAAACCAATAGAAAAATAGTCTCTGGGAATGTGGTCAAGAACCAAGGCCTGA 1679
Db |||||

QY 1647 GTGTCTTTCTCAACCCGTGCAAAAGCCGTGTTCTTCCCGGGAACCCAGGAAAGGATCCGC 1706
Db |||||
QY 1680 GTGTCTTTCTCAACCCGTGCAAAAGCCGTGTTCTTCCCGGGAACCCAGGAAAGGATCCGC 1739
Db |||||
QY 1707 TACTCAAAAAACCAAGAATTTAAAGGAGTTTCTTAAATTTTCGACCTTGTCTTCTGAAGCTCA 1766
Db |||||
QY 1740 TACTCAAAAAACCAAGAATTTAAAGGAGTTTCTTAAATTTTCGACCTTGTCTTCTGAAGCTCA 1799
Db |||||
QY 1767 CTTTTCAGTGCCATTGATGTGAGATGTGCTGGAGTGGCTATTAAACCTTTTTTTCCTAAAG 1826
Db |||||
QY 1800 CTTTTCAGTGCCATTGATGTGAGATGTGCTGGAGTGGCTATTAAACCTTTTTTTCCTAAAG 1859
Db |||||
QY 1827 ATTATTGTTAAATAGATATTGTGGTTTGGGGAAGTTGAATTTTTTATAGGTTAAATGTCA 1886
Db |||||
QY 1860 ATTATTGTTAAATAGATATTGTGGTTTGGGGAAGTTGAATTTTTTATAGGTTAAATGTCA 1919
Db |||||
QY 1887 TTTTAGAGATGGGAGAGGATATATCTGCAGGCAGCTTCAGCCCATGTTGTGAAACTGAT 1946
Db |||||
QY 1920 TTTTAGAGATGGGAGAGGATATATCTGCAGGCAGCTTCAGCCCATGTTGTGAAACTGAT 1979
Db |||||
QY 1947 AAAAGCAACTTAGCAAGGCTCTTTTTCATTATTTTTTATGTTTTCACCTTATAAAGTCTTAG 2006
Db |||||
QY 1980 AAAAGCAACTTAGCAAGGCTCTTTTTCATTATTTTTTATGTTTTCACCTTATAAAGTCTTAG 2039
Db |||||
QY 2007 GTAACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAAAGTACTACTATTGATTA 2066
Db |||||
QY 2040 GTAACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAAAGTACTACTATTGATTA 2099
Db |||||
QY 2067 GAGCCTAACCCAGGTTAACTGCAAGAAGAGGCGGGGATACTTTTCAGCTTCCATGTAAGT 2126
Db |||||
QY 2100 GAGCCTAACCCAGGTTAACTGCAAGAAGAGGCGGGGATACTTTTCAGCTTCCATGTAAGT 2159
Db |||||
QY 2127 TATGCATAAAGCCAAATGTAGTCCAGTTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCA 2186
Db |||||
QY 2160 TATGCATAAAGCCAAATGTAGTCCAGTTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCA 2219
Db |||||
QY 2187 CTTCAATACACACTCATGAACCTCCTGATGGAACAATAAACAGGCCCAAGCCCTGTGGTATGA 2246
Db |||||
QY 2220 CTTCAATACACACTCATGAACCTCCTGATGGAACAATAAACAGGCCCAAGCCCTGTGGTATGA 2279
Db |||||
QY 2247 TGTGCACACTTGTCTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTTG 2306
Db |||||
QY 2280 TGTGCACACTTGTCTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTTG 2339
Db |||||
QY 2307 GTGACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCAATTTATTTCCA 2366
Db |||||
QY 2340 GTGACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCAATTTATTTCCA 2399
Db |||||
QY 2367 TGGACATTTTAGTTAGTGTCTTTTATATACCAGGCATGATGCTGAGTGACACTCTTGTGTA 2426
Db |||||
QY 2400 TGGACATTTTAGTTAGTGTCTTTTATATACCAGGCATGATGCTGAGTGACACTCTTGTGTA 2459
Db |||||
QY 2427 TATTTCCAAATTTTGTATAGTGTGCTGACATATTTGAAATCAAAATATTAAGACTTTTC 2486
Db |||||
QY 2460 TATTTCCAAATTTTGTATAGTGTGCTGACATATTTGAAATC-ATATATTAAGACTTTTC 2518
Db |||||
QY 2487 AAAAATTTGGTCCCTGGTTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTGTTG 2546
Db |||||
QY 2519 AAAGATGAGGTCCCTGGTTTTTTCATGGCAACTTGATCAGTAAGGATTTACCTCTGTGTTG 2578
Db |||||
QY 2547 GAACTAAAAACCATTTACTATATGTTAGACAAGACATTTTTTTTTTTCCTTCCTGAAAAA 2606
Db |||||
QY 2579 TAACTAAAAACCATCTACTATATGTTAGACATGACATTTCTTTTCTCTCTCTCTCTGAAAAA 2638
Db |||||
QY 2607 -AAATGAGGGAAGAGACAAAAAATAAAAAA 2636
Db |||||
QY 2639 TAAAGTGTGGGAAGAGACAAAAAATAAAAAA 2669
Db |||||

```

GENERAL INFORMATION:
APPLICANT: Michele Sanicola-Nadel
          Joseph V. Bonventre
          Catherine A. Hession
          Takaharu Ichimura
          Henry Wei
          Richard L. Cate
TITLE OF INVENTION: MODULATORS OF TISSUE REGENERATION
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Biogen, Inc.
STREET: 14 Cambridge Center
CITY: Cambridge
STATE: MA
COUNTRY: USA
ZIP: 02142
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/197,970B
FILING DATE: 23-NO. 6664385-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/018,228
FILING DATE: 24-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Levine, Leslie M.
REGISTRATION NUMBER: 35,245
REFERENCE/DOCKET NUMBER: A010 PCT CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 679-2810
TELEFAX: (617) 679-2838
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 2303 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 107..1822
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-197-970B-4

```

QY	31	CTTGGTCCGCTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGGATTC	90
Db	78	CTGGCTGAGGCGGGCGCTCCGAGTCAGCATGGAAGTCTCTCGGGGTCTCTGGTATTC	137
QY	91	TGCTCCTGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTCATGATGTCTGGGCA	150
Db	138	TGCTGTGCTGCAGGACTGCCGCTCCAGCGGCCAAAGCGTTCCGTGATGTCTGGGCC	197
QY	151	ATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTAATGCTGGTCTTCTGATG	210
Db	198	ATGAGCAGTATCCGGATCACATGAGGGAGAACCAATTAACGTGGTCTTTCAGATG	257
QY	211	AAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGAAA	270
Db	258	AAATGAATGGGATGAACAGCTGTATCCAGTGTGGAGGAGGGAGAGGCAGATGGAAG	317
QY	271	ACTCCTGGAAGGAGCGCGTGTGCAGGCGGTCTGACCAGTGACTCACCAGCCCTCGTGG	330
Db	318	ACTCCTGGGAAGGAGCGCGTGTGCAGGCAGCCCTAACCACTGATTACCGGCCTTGGTGG	377
QY	331	GCTCAAATATAACATTTCCGGTGAACCTGTATATCCCTAGATGCCAAAAGGAAGATGCCA	390

Db 1458 TGTGCCTCCTGTCCGTGAGGAGAGCCCTTCAATGGGTCCGGCACGTACTGTGAATTCA 1517
QY 1428 CCCTGGGGGATGACACAAGCCCTGGCTCTCACGAGCACCTGTATTCTGTCTCTGACAGAG 1487
Db 1518 CTCTGGGAGACGATGCAAGCCCTGGCCCTCACAGCGCCCTGATCTCTATCCCTGGCAAG 1577
QY 1488 ACCAGCCTCGCCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGCTTGGCCA 1547
Db 1578 ACCTAGGCTCCCTCTGAGAACAGTGAATGGTGTCTGATCTCCATTTGGCTGCCTGGCCA 1637
QY 1548 TATTTGTCACCTGTGATCTCCCTCTTGGTGACAAAAACACAAGGATACAAACCAATAG 1607
Db 1638 TGTGTGTCACCATGGTTACCATCTTGTCTGTACAAAAACACAAGACGTACAAGCCAATAG 1697
QY 1608 AAAATAGTCTCGGAATGTGGTCAGAAAGCAAGCAAGGCTGAGTGTCTTTCTCAACCGTGCAA 1667
Db 1698 GAAACTGCACAGAAAGTGGTCAAGGGCAAGGCTGAGTGTCTTCTCAGCCATGCAA 1757
QY 1668 AAGCCCGTGTCTTCCCGGAAACCAAGGAAAGGATCCGCTACTCAAAAAACCAAGAAATTA 1727
Db 1758 AAGCCCGTGTCTCCCGAGGAGACCGGAGAAGGATCCACTGTCCAGGACAAGCCATGGA 1817
QY 1728 AAGGAGTTCTTAAATTCGACCTTGTCTGAAGCTCACATTTTCAGTGCCATTGATGTG 1787
Db 1818 TGCTCTAAGTCTTCACTCTCACTTCTGACTGGGAACCCCACTCTCTGTG-CATGTATGIG 1876
QY 1788 AGATGTCTGGAG---TGGCTATTAAACCTTTTTCCTAAAGATTATTGTTAAATAGAT 1843
Db 1877 AGCTGTGCAGAAGTACATGACTGTGTAGCTGTGTTTCTACCGGATTATTGTAAATGTAT 1936
QY 1844 ATTGTGGTTTGGGAAGTTGAATTTTATAGGTTA 1879
Db 1937 ATCATGGTTTAGGGAGCGGTAGTTAATTGGCATTTTA 1972

RESULT 6
US-09-383-586-27
; Sequence 27, Application US/09383586
; Patent No. 6242419
; GENERAL INFORMATION:
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Abernethy, Nevin
; APPLICANT: Onrust, Rene
; APPLICANT: Kumble, Anand
; APPLICANT: Murison, Greg
; TITLE OF INVENTION: Compounds isolated from stromal cells
; FILE REFERENCE: 11000.1037cl
; CURRENT APPLICATION NUMBER: US/09/383,586
; CURRENT FILING DATE: 1999-08-26
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 2213
; TYPE: DNA
; ORGANISM: Mouse
US-09-383-586-27

Query Match 39.4%; Score 1047.8; DB 3; Length 2213;
Best Local Similarity 74.1%; Pred. No. 6.7e-216;
Matches 1404; Conservative 0; Mismatches 437; Indels 54; Gaps 4;
QY 40 TCGGTCCGTGAGAATTCAGCATGGAATGTCTCTACTATTTCCTGGGATTTCTGCTCCTGG 99
Db 3 TGCAGGCGCTCGAGTCAGCATGGAAGTCTCTGGGGTCTCTGGGATTTCTGCTGCTGG 62
QY 100 CTGCAAGATTGCCACTTGTATGCCGCCAAACGATTTTCATGATGTCTGGCAATGAAAGAC 159
Db 63 CTGCAGGACTGCTCTCCAGGCTGCCAAGCGATTTCTGTGATGTGTGGGCCATGAACAGT 122
QY 160 CTTCTGCTTACATGAGGGAGCACAATCAATTAATGGCTGGTCTTCTGATGAAAATGACT 219

Db 123 ATCCCAATCATGATGAGAGAGCAACAACCAATTACGTGGTGGTCTTTCGGATGAAAAATGAAT 182
QY 220 GGAATGAAAAAATCTTACCCAGTGTGGAGCGGGGAGACATGAGGTGGAATAAACTCCTGGA 279
Db 183 GGGATGAACAACCTGTATCCAGTGTGGAGAGGGGAGACGGCAGGTGGAAGGACTCCTGGG 242
QY 280 AGGAGGCCGTGTGCAGGCGGTCTCTGACAGTGAATCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACA 339
Db 243 AAGGAGGCCGTGTGCAGGCAGTCTCTGACAGTGAATCCCGGCTCTGGTGGGTCCAATA 302
QY 340 TAAACATTTGCGGTGAACCTGATATTCCTCTAGATGCCAAAAGGAAGATGCCAATGGCAACA 399
Db 303 TCACCTTTGTGGTGAACCTGGTGTTCCTGAGTGAATGGGCTCTGATGTTTACA 459
QY 400 TAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACA 459
Db 363 TCGTCTATGAGAAGAACTGCAGGAATGATTTGGGACTGACCTCTGACCTGCATGCTACA 422
QY 460 ACTGACAGCATGGTTCAGAGGACAGTGAACGGGGAATAATGGCAACCGGCCAAAGCCATCATA 519
Db 423 ACTGACTGCAGGGGCAGATGATGGTGAAGTGGCAAGTGGCAAGCCAGCATC 482
QY 520 ACCTCTTCCCTGATGGGAAACCTTTTCTCACACCCCGGATGGAGAAGATGGAATTTCA 579
Db 483 TCAGTTCCCGACAGGAGGCCCTTCCCTCGCCCCCTGATGGAGAATAATGGAGCTTTG 542
QY 580 TCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGATGTTTCAGTGAGAG 639
Db 543 TCTACGTCTTTCACACACTTGGCCAGTATTTCCAAAAAACTGGGTGGTTCAGCACGGG 602
QY 640 TTTCTGTGAACACAGCCCAATGTGACACTTGGGCCCTCAACTCATGGAAGTGAATGCTTACA 699
Db 603 TTTCTATAAACACAGTCAACTTGACAGCTGGCCCTCAGGTTCATGGAAGTGAATGCTTCTTC 662
QY 700 GAAGACATGGACGGGCATATGTTTCCCATCGCACAAAGTGAAGATGTGTACGTGTAACAG 759
Db 663 GAAGATACGGCCGGGCATACATTCCCATCTCGAAGGTGAAGATGTGTATGTATAACAG 722
QY 760 ATCAGATTCCTGTGTGTGACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAA 819
Db 723 ATCAGATCCCTGTATTTCGTGACCATGTCCCAAGAGAATGACAGGAACCTTGTCTGATGAGA 782
QY 820 CCTTCC-CAAAGATCTCCCATTAATGTTTTGATGTCTCTGATTCATGATCCTAGCCACTTCC 878
Db 783 TCTTCCCTCAGAGACCTCCCATCGTCTTCGATGTCTCTCATTCATGATCCAGCCACTTCC 842
QY 879 TCAATTATTCTACCATTAACATAAGTGGAGCTTCGGGGATAATACTGGCCTGTGTTGTTT 938
Db 843 TCAACGACTCTGCCATTTCTTACAAGTGGAACTTTGGGGAACAACACTGGCCTGTGTTGCT 902
QY 939 CCACCAATCATACTGTGAATCACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCA 998
Db 903 CCAACAATCACACTTTGAATCACACTTATGTCTCAATGGAACCTTCAACCTTAACCTCA 962
QY 999 CTGTGAAGCTGCAGCACCCAGGACCTTG----- 1026
Db 963 CCGTGCAAACTGCAGTGCCCGGGCCATGSCCTTCCCTTTCGCTTCGACTCCGCTCCAC 1022
QY 1027 -----TCCGCCACCGCCACCCACCACCCAGACCTTCAAAACCCACCCCTT 1070
Db 1023 CTTCAACTCCGCCCTCACCTCCGCCCTCACCTCTGCCCCACATTATCAACACCTAGCCCT 1082
QY 1071 CTTTAGGACCTGTGTTGAGACAACCCCTTGAGCTGAGTAGGATTCCTGATGAAAACTGCC 1130
Db 1083 CTTTAAATGCCTACTGTTTACAATCCATCCATGGAGCTGAGTGACATTTCCAATGAAAACTGCC 1142
QY 1131 AGATTAAACAGATATGGCCACTTTCAAGCCACCACCATCAAAATTTAGAGGGGAATCTTAGAG 1190
Db 1143 GAATAAACAGATATGGTACTCTCAGAGCCACCACCATCAAAATTTAGAGGGGATCCTTGAAG 1202
QY 1191 TTAACATCATCCAGATGACAGAGCTCTGATGCCGTTGCCATGGCTGAAAGCTCCCTAA 1250
Db 1203 TCAGCATCATGCAGATAGCAGATGTCCCCTATGCCACACCGCAGCCTGCCAACTCCCTGA 1262


```

; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 105
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(619)
; OTHER INFORMATION: n = A,T,C or G
US-09-643-597-105

Query Match      20.5%; Score 545.6; DB 4; Length 619;
Best Local Similarity 93.2%; Pred. No. 3.2e-108;
Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 2010 ACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAAAGCTACTATTGATTAGAG 2069
Db 2 ACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAAAGCTACTATTGATTAGAG 61

QY 2070 CCTAACCCAGGTTAACTGCAAGAAAGAGCGGGGATCTTTCCAGCTTTCCATGTAACCTGTAT 2129
Db 62 CCTAACCCAGGTTAACTGCAAGAAAGAGCGGGGATCTTTCCAGCTTTCCATGTAACCTGTAT 121

QY 2130 GCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTTCCAAGCTAACTGAATCCCACCTT 2189
Db 122 GCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTTCCAAGCTAACTGAATCCCACCTT 181

QY 2190 CAATACACACTCATGAACCTCCTGATGGAAACAATAACAGGCCCAAGCCTGTGGTATGATGT 2249
Db 182 CAATACACACTCATGAACCTCCTGATGGAAACAATAACAGGCCCAAGCCTGTGGTATGATGT 241

QY 2250 GCACACTTGTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTGGTG 2309
Db 242 GCACACTTGTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTGGTG 301

QY 2310 ACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATCCATGG 2369
Db 302 ACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATCCATGG 361

QY 2370 ACATTAGTGTAGTGCTTTTATATACCAAGGCATGATGCTGAGTGACACTCTTGTGTATAT 2429
Db 362 ACATTAGTGTAGTGCTTTTATATACCAAGGCATGATGCTGAGTGACACTCTTGTGTATAT 421

QY 2430 TTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAATATTAAGACTTTCCTCAA 2489
Db 422 TTCCAAATTTTGTACAGTCGCTGCACATATTTGAAATC-ATATATTAAGACTTTCCTCAA 480

QY 2490 AATTGGTCCCTGGTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTGGAA 2549
Db 481 AATGAGTCCCTGGTTTTTCATGGCAACTTGATCAGTAAGGATTCNCCTCTGTGGTA 540

QY 2550 CTAAACCATTTACTATATGTTAGCAAGACATTTTTTTTTTTTCTCTCTGAAAAAAA 2609
Db 541 CTTAAACATCTACTATATNGTTNANATGAAATTCCTTTTCCCNCCCTCCCGAAAAANA 600

QY 2610 ATGAGGGAAGAGACAAAAA 2628
Db 601 AAGTGGTGGGAAAAAAA 619
```

```

RESULT 9
US-09-643-597-121
; Sequence 121, Application US/09643597
; Patent No. 6426072
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C11
; CURRENT APPLICATION NUMBER: US/09/643,597
; CURRENT FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 369
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 121
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(619)
; OTHER INFORMATION: n = A,T,C or G
US-09-643-597-121
```

```

Query Match      20.5%; Score 545.6; DB 4; Length 619;
Best Local Similarity 93.2%; Pred. No. 3.2e-108;
Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 2010 ACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAAAGCTACTATTGATTAGAG 2069
Db 2 ACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAAAGCTACTATTGATTAGAG 61

QY 2070 CCTAACCCAGGTTAACTGCAAGAAAGAGCGGGGATCTTTCCAGCTTTCCATGTAACCTGTAT 2129
Db 62 CCTAACCCAGGTTAACTGCAAGAAAGAGCGGGGATCTTTCCAGCTTTCCATGTAACCTGTAT 121

QY 2130 GCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTTCCAAGCTAACTGAATCCCACCTT 2189
Db 122 GCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTTCCAAGCTAACTGAATCCCACCTT 181

QY 2190 CAATACACACTCATGAACCTCCTGATGGAAACAATAACAGGCCCAAGCCTGTGGTATGATGT 2249
Db 182 CAATACACACTCATGAACCTCCTGATGGAAACAATAACAGGCCCAAGCCTGTGGTATGATGT 241

QY 2250 GCACACTTGTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTGGTG 2309
Db 242 GCACACTTGTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTGGTG 301

QY 2310 ACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATCCATGG 2369
Db 302 ACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATCCATGG 361

QY 2370 ACATTAGTGTAGTGCTTTTATATACCAAGGCATGATGCTGAGTGACACTCTTGTGTATAT 2429
Db 362 ACATTAGTGTAGTGCTTTTATATACCAAGGCATGATGCTGAGTGACACTCTTGTGTATAT 421

QY 2430 TTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAATATTAAGACTTTCCTCAA 2489
Db 422 TTCCAAATTTTGTACAGTCGCTGCACATATTTGAAATC-ATATATTAAGACTTTCCTCAA 480

QY 2490 AATTGGTCCCTGGTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTGGAA 2549
Db 481 AATGAGTCCCTGGTTTTTCATGGCAACTTGATCAGTAAGGATTCNCCTCTGTGGTA 540
```


QY 2550 CTAAAAACCAATTTACTATATGTTAGACAAGACATTTTTTTTTTCCCTCCTGAAAAAAA 2609
Db 541 CTTAAAAACATCTACTATATNGTTNANATGAAATTCCTTTTCCCNCCCTCCCGAAAAANA 600
QY 2610 ATGAGGGAAGAGACAAAAA 2628
Db 601 AAGTGGTGGGAAAAAAA 619

RESULT 10
US-09-480-884A-105
; Sequence 105, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 105
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(619)
; OTHER INFORMATION: n = A,T,C or G
US-09-480-884A-105

Query Match 20.5%; Score 545.6; DB 4; Length 619;
Best Local Similarity 93.2%; Pred. No. 3.2e-108;
Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
QY 2010 ACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 2069
Db 2 ACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 61
QY 2070 CCTAACCCAGGTTAACTGCAAGAAGAGGCGGGATCTTTCAGCTTCCATCTAACTGTAT 2129
Db 62 CCTAACCCAGGTTAACTGCAAGAAGAGGCGGGATCTTTCAGCTTCCATCTAACTGTAT 121
QY 2130 GCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCTCAAGCTAACTGAATCCCAC 2189
Db 122 GCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCTCAAGCTAACTGAATCCCAC 181
QY 2190 CAATACACACTCATGAATCCTGTATGGAAACAATAACAGGCCCCAAGCCTGTGGTATGATGT 2249
Db 182 CAATACACACTCATGAATCCTGTATGGAAACAATAACAGGCCCCAAGCCTGTGGTATGATGT 241
QY 2250 GCACACTTGCTAGACTCAGAAAAAATACTACTCTATAAATGGGTGGGAGTATTTGGTG 2309
Db 242 GCACACTTGCTAGACTCAGAAAAAATACTACTCTATAAATGGGTGGGAGTATTTGGTG 301
QY 2310 ACAACCTACTTTTGGCTGAGTGAAGGAATGATATTCATATATATTCATTTATTCATGG 2369
Db 302 ACAACCTACTTTTGGCTGAGTGAAGGAATGATATTCATATATATTCATTTATTCATGG 361
QY 2370 ACATTTAGTTAGTGCTTTTATATACCGGCATGATGCTGAGTGACACTCTTTGTGTATAT 2429
Db 362 ACATTTAGTTAGTGCTTTTATATACCGGCATGATGCTGAGTGACACTCTTTGTGTATAT 421
QY 2430 TTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCATAATTAAGACTTTTCCAAA 2489
Db 422 TTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATC-ATATATTAAGACTTTCCAAA 480

QY 2490 AATTTGGTCCCTGGTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTCTTTGGAA 2549
Db 481 AATGAAGTCCCTGGTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTCTTTGGTA 540
QY 2550 CTAAAAACCAATTTACTATATGTTAGACAAGACATTTTTTTTTTCCCTCCTGAAAAAAA 2609
Db 541 CTTAAAAACATCTACTATATNGTTNANATGAAATTCCTTTTCCCNCCCTCCCGAAAAANA 600
QY 2610 ATGAGGGAAGAGACAAAAA 2628
Db 601 AAGTGGTGGGAAAAAAA 619

RESULT 11
US-09-480-884A-121
; Sequence 121, Application US/09480884A
; Patent No. 6482597
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C6
; CURRENT APPLICATION NUMBER: US/09/480,884A
; CURRENT FILING DATE: 2001-08-27
; NUMBER OF SEQ ID NOS: 330
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 121
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(619)
; OTHER INFORMATION: n = A,T,C or G
US-09-480-884A-121

Query Match 20.5%; Score 545.6; DB 4; Length 619;
Best Local Similarity 93.2%; Pred. No. 3.2e-108;
Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
QY 2010 ACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 2069
Db 2 ACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 61
QY 2070 CCTAACCCAGGTTAACTGCAAGAAGAGGCGGGATCTTTCAGCTTCCATCTAACTGTAT 2129
Db 62 CCTAACCCAGGTTAACTGCAAGAAGAGGCGGGATCTTTCAGCTTCCATCTAACTGTAT 121
QY 2130 GCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCTCAAGCTAACTGAATCCCAC 2189
Db 122 GCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCTCAAGCTAACTGAATCCCAC 181
QY 2190 CAATACACACTCATGAATCCTGTATGGAAACAATAACAGGCCCCAAGCCTGTGGTATGATGT 2249
Db 182 CAATACACACTCATGAATCCTGTATGGAAACAATAACAGGCCCCAAGCCTGTGGTATGATGT 241
QY 2250 GCACACTTGCTAGACTCAGAAAAAATACTACTCTATAAATGGGTGGGAGTATTTGGTG 2309
Db 242 GCACACTTGCTAGACTCAGAAAAAATACTACTCTATAAATGGGTGGGAGTATTTGGTG 301
QY 2310 ACAACCTACTTTTGGCTGAGTGAAGGAATGATATTCATATATATTCATTTATTCATGG 2369
Db 302 ACAACCTACTTTTGGCTGAGTGAAGGAATGATATTCATATATATTCATTTATTCATGG 361
QY 2370 ACATTTAGTTAGTGCTTTTATATACCGGCATGATGCTGAGTGACACTCTTTGTGTATAT 2429
Db 362 ACATTTAGTTAGTGCTTTTATATACCGGCATGATGCTGAGTGACACTCTTTGTGTATAT 421
QY 2430 TTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCATAATTAAGACTTTTCCAAA 2489
Db 422 TTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCATAATTAAGACTTTTCCAAA 2489

Db 422 TTCCAAATTTTGTACAGTCGCTGCACATATTTTGAATC-ATATATTAAGACTTCCAAA 480
QY 2490 AATTGGTCCTGGTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTGGAA 2549
Db 481 AATGAAGTCCTGGTTTTCATGGCAACTTGATCAGTAAGGATTCNCCTCTGTTGGTA 540
QY 2550 CTAAACCATTTACTATATGTTAGACAAGACATTTTTCCTTCCCTGAAAAAAA 2609
Db 541 CTAAACATCTACTATATNGTTNANATGAAATTCCTTTCCCNCTCCCGAAAAANA 600
QY 2610 ATGAGGAAGAGACAAAA 2628
Db 601 AAGTGGTGGGAAAAANA 619

RESULT 12
US-09-542-615A-105
; Sequence 105, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 105
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(619)
; OTHER INFORMATION: n = A,T,C or G
US-09-542-615A-105

Query Match 20.5%; Score 545.6; DB 4; Length 619;
Best Local Similarity 93.2%; Pred. No. 3.2e-108;
Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 2010 ACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 2069
Db 2 ACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 61
QY 2070 CCTAACCCAGGTTAACTGCAAGAGAGGGGGGATACTTTCAGCTTTCCTGTAAGTGTAT 2129
Db 62 CCTAACCCAGGTTAACTGCAAGAGAGGGGGGATACTTTCAGCTTTCCTGTAAGTGTAT 121
QY 2130 GCATAAGCCAAATGATGTCAGTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCACCT 2189
Db 122 GCATAAGCCAAATGATGTCAGTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCACCT 181
QY 2190 CAATACACTCATGAATCCTGTATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGT 2249
Db 182 CAATACACTCATGAATCCTGTATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGT 241
QY 2250 GCACACTTGCTAGACTCAGAAAAATACTACTCTCATAAATGGGTGGAGTATTTGGTG 2309
Db 242 GCACACTTGCTAGACTCAGAAAAATACTACTCTCATAAATGGGTGGAGTATTTGGTG 301
QY 2310 ACAACCTACTTTGCTTGCTGAGTGAAGGAATGATATTCATATATTCATTTATCCATGG 2369
Db 302 ACAACCTACTTTGCTTGCTGAGTGAAGGAATGATATTCATATATTCATTTATCCATGG 361
QY 2370 ACATTTAGTAGTGCTTTTATATACAGGCATGATGCTGAGTGACACTCTTGTGTATAT 2429

Db 362 ACATTTAGTTAGTCTTTTATATACCAGGCATGATGCTGAGTGACACTCTTGTGTATAT 421
QY 2430 TTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAAAATATTAAGACTTCCAAA 2489
Db 422 TTCCAAATTTTGTACAGTCGCTGCACATATTTGAAATC-ATATATTAAGACTTCCAAA 480
QY 2490 AATTGGTCCTGGTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTGGAA 2549
Db 481 AATGAAGTCCTGGTTTTCATGGCAACTTGATCAGTAAGGATTCNCCTCTGTTGGTA 540
QY 2550 CTAAACCATTTACTATATGTTAGACAAGACATTTTTCCTTCCCTGAAAAAAA 2609
Db 541 CTAAACATCTACTATATNGTTNANATGAAATTCCTTTCCCNCTCCCGAAAAANA 600
QY 2610 ATGAGGAAGAGACAAAA 2628
Db 601 AAGTGGTGGGAAAAANA 619

RESULT 13
US-09-542-615A-121
; Sequence 121, Application US/09542615A
; Patent No. 6518256
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542,615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 121
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(619)
; OTHER INFORMATION: n = A,T,C or G
US-09-542-615A-121

Query Match 20.5%; Score 545.6; DB 4; Length 619;
Best Local Similarity 93.2%; Pred. No. 3.2e-108;
Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 2010 ACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 2069
Db 2 ACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 61
QY 2070 CCTAACCCAGGTTAACTGCAAGAGAGGGGGGATACTTTCAGCTTTCCTGTAAGTGTAT 2129
Db 62 CCTAACCCAGGTTAACTGCAAGAGAGGGGGGATACTTTCAGCTTTCCTGTAAGTGTAT 121
QY 2130 GCATAAGCCAAATGATGTCAGTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCACCT 2189
Db 122 GCATAAGCCAAATGATGTCAGTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCACCT 181
QY 2190 CAATACACTCATGAATCCTGTATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGT 2249
Db 182 CAATACACTCATGAATCCTGTATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGT 241
QY 2250 GCACACTTGCTAGACTCAGAAAAATACTACTCTCATAAATGGGTGGAGTATTTGGTG 2309
Db 242 GCACACTTGCTAGACTCAGAAAAATACTACTCTCATAAATGGGTGGAGTATTTGGTG 301
QY 2310 ACAACCTACTTTGCTTGCTGAGTGAAGGAATGATATTCATATATTCATTTATCCATGG 2369

Db 302 ACAACCTACTTTGGCTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATTCATGG 361
QY 2370 ACATTTAGTTAGTGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTTGTGTATAT 2429
Db 362 ACATTTAGTTAGTGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTTGTGTATAT 421
QY 2430 TTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAAATATTAAGACTTTCCAAA 2489
Db 422 TTCCAAATTTTGTACAGTCGCTGCACATATTTGAAATC-ATATATTAAGACTTTCCAAA 480
QY 2490 AATTTGGTCCCTGGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTTGTGTATAT 2549
Db 481 AATGAAGTCCCTGGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTTGTGTATAT 540
QY 2550 CTAAACCAATTTACTATATATGTTAGACAGACATTTTCTTTTCTTCTGAAAAAAA 2609
Db 541 CTTAAACATCTACTATATATNGTTNANATGAAATTCCTTTTCCCNCCCTCCCGAAAAANA 600
QY 2610 ATGAGGGAAGAGACAAAAA 2628
Db 601 AAGTGGTGGGAAAAAAA 619

RESULT 14
US-09-606-421B-105
; Sequence 105, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 105
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(619)
; OTHER INFORMATION: n = A,T,C or G
US-09-606-421B-105

Query Match 20.5%; Score 545.6; DB 4; Length 619;
Best Local Similarity 93.2%; Pred. No. 3.2e-108;
Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
QY 2010 ACTAGTAGGATAGAAACACTGTGTCCTGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 2069
Db 2 ACTAGTAGGATAGAAACACTGTGTCCTGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 61
QY 2070 CCTAACCCAGGTTAACTGCAAGAGAGCGGGGATACCTTTCCAGCTTCCATGTAAGTGTAT 2129
Db 62 CCTAACCCAGGTTAACTGCAAGAGAGCGGGGATACCTTTCCAGCTTCCATGTAAGTGTAT 121
QY 2130 GCATAAGCCAAATGTAGTCCAGTTTCTAAGATCATGTTCCAGCTAACTGAATCCCACTT 2189
Db 122 GCATAAGCCAAATGTAGTCCAGTTTCTAAGATCATGTTCCAGCTAACTGAATCCCACTT 181
QY 2190 CAATACACACTCATGAACTCCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGT 2249

Db 182 CAATACACACTCATGAACCTCCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGT 241
QY 2250 GCACACTTGTAGACTCAGAAAAAATACTACTCTCATAAATGGTGGAGTATTTTGGTG 2309
Db 242 GCACACTTGTAGACTCAGAAAAAATACTACTCTCATAAATGGTGGAGTATTTTGGTG 301
QY 2310 ACAACCTACTTTGGCTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATTCATGG 2369
Db 302 ACAACCTACTTTGGCTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATTCATGG 361
QY 2370 ACATTTAGTTAGTGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTTGTGTATAT 2429
Db 362 ACATTTAGTTAGTGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTTGTGTATAT 421
QY 2430 TTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAAATATTAAGACTTTCCAAA 2489
Db 422 TTCCAAATTTTGTACAGTCGCTGCACATATTTGAAATC-ATATATTAAGACTTTCCAAA 480
QY 2490 AATTTGGTCCCTGGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTTGTGTATAT 2549
Db 481 AATGAAGTCCCTGGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTTGTGTATAT 540
QY 2550 CTAAACCAATTTACTATATATGTTAGACAGACATTTTCTTTTCTTCTGAAAAAAA 2609
Db 541 CTTAAACATCTACTATATATNGTTNANATGAAATTCCTTTTCCCNCCCTCCCGAAAAANA 600
QY 2610 ATGAGGGAAGAGACAAAAA 2628
Db 601 AAGTGGTGGGAAAAAAA 619

RESULT 15
US-09-606-421B-121
; Sequence 121, Application US/09606421B
; Patent No. 6531315
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C9
; CURRENT APPLICATION NUMBER: US/09/606,421B
; CURRENT FILING DATE: 2000-06-28
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 121
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(619)
; OTHER INFORMATION: n = A,T,C or G
US-09-606-421B-121
Query Match 20.5%; Score 545.6; DB 4; Length 619;
Best Local Similarity 93.2%; Pred. No. 3.2e-108;
Matches 577; Conservative 0; Mismatches 41; Indels 1; Gaps 1;
QY 2010 ACTAGTAGGATAGAAACACTGTGTCCTGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 2069
Db 2 ACTAGTAGGATAGAAACACTGTGTCCTGAGAGTAAGGAGAGAGCTACTATTGATTAGAG 61
QY 2070 CCTAACCCAGGTTAACTGCAAGAGAGCGGGGATACCTTTCCAGCTTCCATGTAAGTGTAT 2129
Db 62 CCTAACCCAGGTTAACTGCAAGAGAGCGGGGATACCTTTCCAGCTTCCATGTAAGTGTAT 121

QY	2130	GCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCACIT	2189
Db	122	GCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCACIT	181
QY	2190	CAATACACACTCATGAACCTCCTGATGGGAACAATAACAGGCCCAAGCCTGTGGTATGATGT	2249
Db	182	CAATACACACTCATGAACCTCCTGATGGGAACAATAACAGGCCCAAGCCTGTGGTATGATGT	241
QY	2250	GCACACTTGCCTAGACTCAGAAAAAATACTACTCTCATAAATGGTGGGAGTATTTGGTG	2309
Db	242	GCACACTTGCCTAGACTCAGAAAAAATACTACTCTCATAAATGGTGGGAGTATTTGGTG	301
QY	2310	ACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTCCATGG	2369
Db	302	ACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTCCATGG	361
QY	2370	ACATTAGTTAGTGTCTTTTATATACCAGGCATGATGCTGAGTGACACTCTTGTGTATAT	2429
Db	362	ACATTAGTTAGTGTCTTTTATATACCAGGCATGATGCTGAGTGACACTCTTGTGTATAT	421
QY	2430	TTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAAATATTAAAGACTTTCCAAA	2489
Db	422	TTCCAAATTTTGTACAGTCGCTGCACATATTTGAAATC-ATATATTAAAGACTTTCCAAA	480
QY	2490	AATTGGTCCCTGGTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTGTTGGAA	2549
Db	481	AATGAAGTCCCTGGTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTGTTGGTA	540
QY	2550	CTAAAACCATTTACTATATGTTAGACAAGACATTTTTCCTTTCCTGAAAAAAA	2609
Db	541	CTTAAAAACATCTACTATATNGTTNANATGAAATTCCTTTCCCCNCCTCCCGAAAAANA	600
QY	2610	ATGAGGGAAGAGACAAAAA	2628
Db	601	AAGTGGTGGGAAAAAAA	619

Search completed: September 10, 2004, 10:01:25
Job time : 218 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 10, 2004, 09:57:48 ; Search time 200 Seconds
(without alignments)
7518.589 Million cell updates/sec

Title: US-10-039-272-1
Perfect score: 4850
Sequence: 1 cggcacgagggcccgagga.....aaaaaaaaaaaaaaaa 2661

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10039272/runat_07092004_144222_21609/app_query.fasta_1.2823
-DB=A_Geneseq_29Jan04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10039272_@CGN_1_1_265_@runat_07092004_144222_21609 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	2999	61.8	560 2 AAW35382 Murine me
2	2999	61.8	560 3 AAB11329 Human lun
3	2999	61.8	560 5 ABB78200 Amino aci
4	2999	61.8	560 5 ABB74961 Human lun
5	2999	61.8	560 5 ABP61881 Human lun
6	2999	61.8	560 6 ABU56592 Lung canc
7	2999	61.8	560 6 ABG72962 Human ost
8	2999	61.8	560 6 ABU70852 Human adi
9	2999	61.8	560 7 ADA28315 Human lun
10	2980	61.4	572 7 ADD78235 Human CGD

11	2979	61.4	572 5 AAU83612 Human PRO
12	2979	61.4	572 6 ABU80759 Human PRO
13	2979	61.4	572 6 ABO33725 Novel hum
14	2979	61.4	572 6 ABU82068 Novel hum
15	2979	61.4	572 6 ABJ72248 Human PRO
16	2979	61.4	572 6 ABJ72376 Human PRO
17	2979	61.4	572 6 ABO34271 Human sec
18	2979	61.4	572 7 ABJ72078 Human mem
19	2979	61.4	572 7 ADB83532 Novel hum
20	2979	61.4	572 7 ADB80638 Novel hum
21	2979	61.4	572 7 ADB73179 Novel hum
22	2979	61.4	572 7 ADB78261 Novel hum
23	2979	61.4	572 7 ADB84909 Human PRO
24	2979	61.4	572 7 ADB78015 Novel hum
25	2979	61.4	572 7 ADB87081 Human PRO
26	2979	61.4	572 7 ADB84663 Human PRO
27	2979	61.4	572 7 ADB83778 Novel hum
28	2979	61.4	572 7 ADB72933 Novel hum
29	2979	61.4	572 7 ADC36771 Human PRO
30	2979	61.4	572 7 ADC21761 Human PRO
31	2979	61.4	572 7 ADC49792 Novel hum
32	2979	61.4	572 7 ADC48991 Novel hum
33	2979	61.4	572 7 ADC49508 Novel hum
34	2979	61.4	572 7 ADC47369 Novel hum
35	2979	61.4	572 7 ADC47114 Novel hum
36	2979	61.4	572 7 ADC77989 Novel hum
37	2979	61.4	572 7 ADD06224 Novel hum
38	2979	61.4	572 7 ADC77743 Novel hum
39	2979	61.4	572 7 ADD50706 Novel hum
40	2979	61.4	572 7 ADD50952 Novel hum
41	2979	61.4	572 7 ADD50433 Human PRO
42	2979	61.4	572 7 ADD50187 Human PRO
43	2979	61.4	572 7 ADD51198 Novel hum
44	2979	61.4	572 8 ADC48745 Novel hum
45	2979	61.4	572 8 ADE20916 Novel hum

ALIGNMENTS

RESULT 1
AAW35382
ID AAW35382 standard; protein; 560 AA.
XX
AC AAW35382;
XX
DT 26-FEB-1998 (first entry)
XX
DE Murine metastatic nucleic acid sequence product.
XX
KW Mouse; murine; tumour; cancer; metastatic sequence; detection; diagnosis;
XX treatment; metastasis; hyperplasia; dysplasia; hypertrophy; screening.
OS Mus musculus.
XX
PN WO9718454-A2.
XX
PD 22-MAY-1997.
XX
PF 15-NOV-1996; 96WO-US018567.
XX
PR 16-NOV-1995; 95US-0006838P.
PR 30-JAN-1996; 96US-00594031.
XX
PA (THOM/) THOMPSON T.
XX
PI Thompson T;
XX
DR WPI; 1997-289397/26.
XX

Identifying tumour metastatic sequences - by introducing transfected cells into host mammal and analysing primary and metastatic sequences by differential display PCR.

PS Disclosure; Fig 12CI; 102pp; English.

XX Mouse Urogenital Sinus (UGS) tissue was isolated from 17 day old mouse embryos. The UGS cells were infected with retroviruses, cultured and implanted under the renal capsule of mice. Reconstitutions were harvested 5 weeks later, when they showed signs of distress from the tumour burden. Metastasised tumours were isolated from a site outside the renal capsule. RNA was isolated from primary tumours and metastases, reverse transcribed and subjected to differential display PCR. The sequences were analysed to obtain metastatic sequences, e.g. the sequence encoding the present sequence. The method can be used to detect, diagnose and treat disorders related to metastasis, or treat malignant or non-malignant disorders, e.g. hyperplasia, dysplasia and hypertrophy. The metastatic sequence can be used to screen a biological sample for metastasis, and it or its expression product may also be used to treat a metastatic disorder

XX Sequence 560 AA;

SQ Alignment Scores:

Pred. No.: 4.82e-295 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 2 Gaps: 0

US-10-039-272-1 (1-2661) x AAW35382 (1-560)

QY	60	ATGGAATGTCTCTACTATTTCTCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT	119
Db	1	MetGluCysLeuTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp	20
QY	120	GCCGCCAAACGATTTCATGATGTCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG	179
Db	21	AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40
QY	180	CACAATCAATTAATGGCTGCTTCTGATGAAATGACTGGAATGAAAACTCTACCCA	239
Db	41	HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60
QY	240	GTGTGGAGCGGGAGACATGAGGTGGAAAAAACTCCTGGAGGGAGCGCGTGTGCAGGCG	299
Db	61	ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla	80
QY	300	GTCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTG	359
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100
QY	360	ATATTCCTTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC	419
Db	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120
QY	420	AGAAATGAGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG	479
Db	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140
QY	480	GACAGTACGGGAAAAATGGCACCGGCAAAAGCCATCATACGTTCTCCCTGATGGGAAA	539
Db	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys	160
QY	540	CCTTTTCTCACCACCCCGGATGGAGAAGATGGAATTTTCATCTACGCTTCCACACACTT	599
Db	161	ProPheProHisHisProGlyTrpArgTrpAsnPheIleTyrValPheHisThrLeu	180
QY	600	GGTCAGTATTTCCAGAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT	659
Db	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200
QY	660	GTGACACTTGGGCTCAACTCATGGAAGTCACTGTCTACAGAAGACATGGACGGGCATAT	719
Db	201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220
QY	720	GTTCCCATCGCACAAAGTGAAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTGTG	779

Db	221	ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal	240
QY	780	ACTATGTTCCAGAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC	838
Db	241	ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro	260
QY	839	ATTATGTTTGATGCTCTGATTCATGATCCTAGCCACTTCTCAATTATTCTACCATTAAC	898
Db	261	IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280
QY	899	TACAAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATCTGTGAAT	958
Db	281	TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300
QY	959	CACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCA	1018
Db	301	HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro	320
QY	1019	GGACCTTGTCGCCACCACCCACCCAGACCTTCAAAACCCACCCCTTCTTTAGGA	1078
Db	321	GlyProCysProProProProProProProProProProProProProProProSerLeuGly	340
QY	1079	CCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCTCTGATGAAAACTGCCAGATTAAAC	1138
Db	341	ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn	360
QY	1139	AGATATGGCCACTTTTCAAGCCACCATCACAAATTTAGAGGAATCTTAGAGGTTAACATC	1198
Db	361	ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle	380
QY	1199	ATCCAGATGACAGACGTCCTGATGCCGGTGCCATGGCCCTGAAAGCTCCCTAATAGACTTT	1258
Db	381	IleGlnMetThrAspValLeuMetProValProTyrProGluSerSerLeuIleAspPhe	400
QY	1259	GTCTGACCTGCCAAGGAGGACATTCACCGAGGTCTGTACCATCATTTCTGACCCACC	1318
Db	401	ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr	420
QY	1319	TGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTGCTG	1378
Db	421	CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu	440
QY	1379	ACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCTGGGGGAT	1438
Db	441	ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp	460
QY	1439	GACAAAGCCTGGCTCTCACGACACCCCTGATTTCTGTTCCTGACAGAGACCCAGCCTCG	1498
Db	461	AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer	480
QY	1499	CCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACT	1558
Db	481	ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr	500
QY	1559	GTGATCTCCCTCTTGGTGTAACAAAAACACAAGGAATACAACCAATAGAAAATAGTCCT	1618
Db	501	ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro	520
QY	1619	GGGAATGTGGTCAGAAAGCAAGCGCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTC	1678
Db	521	GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe	540
QY	1679	TTCCCGGAAACCCAGGAAAAGGATCCGCTACTCAAAAACCAAGAATTTAAAGGAGTTTCT	1738
Db	541	PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer	560

RESULT 2
AAB11329
ID AAB11329 standard; protein; 560 AA.
XX
AC AAB11329;
XX

DT 21-FEB-2001 (first entry)
XX Human lung cancer-associated protein L528S.
DE
XX
XX
KW Lung cancer; therapy; treatment; human; tumor; immunogenic; cytostatic;
KW vaccine; detection.
XX
XX Homo sapiens.
OS
XX WO200061612-A2.
PN
XX 19-OCT-2000.
PD
XX
XX 03-APR-2000; 2000WO-US008896.
PF
XX
XX 02-APR-1999; 99US-00285479.
PR
XX 17-DEC-1999; 99US-00466396.
PR
XX 30-DEC-1999; 99US-00476496.
PR
XX 10-JAN-2000; 2000US-00480884.
PR
XX 22-FEB-2000; 2000US-00510376.
XX
XX (CORI-) CORIXA CORP.
PA
XX Wang T, Fan L;
PI
XX WPI; 2000-628399/60.
XX
XX Isolated polypeptide comprising an immunogenic portion of a lung tumor
XX protein is used for detecting and monitoring progression of lung cancer
XX in a patient.
XX
XX Claim 3; Page 203-204; 261pp; English.
XX
XX This invention describes a novel isolated polypeptide (I) which
XX comprising an immunogenic portion of a lung tumor protein or variant (P2)
XX which have cytostatic activity. The polypeptides and polynucleotides are
XX used in compositions and vaccines to inhibit the development of cancer,
XX especially lung cancer, in a patient. Methods described in the invention
XX can be used to monitor the progression of a cancer by carrying out the
XX detection at subsequent time points and comparing the results from the
XX different time points. CD4+ and/or CD8+ T-Cells isolated from a patient
XX are treated with P2, polynucleotides encoding P2 or antigen presenting
XX cells expressing P2 and then administered to the patient to inhibit
XX development of cancer
XX
SQ Sequence 560 AA;

Alignment Scores:
Pred. No.: 4.82e-295 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 3 Gaps: 0

US-10-039-272-1 (1-2661) x AAB11329 (1-560)

QY 60 ATGGAATGCTCTACTATTTCTCTGGGATTTCTGGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db
1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20

QY 120 GCCGCCAAACGATTTCATGATGTCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db
21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40

QY 180 CACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTCTACCCA 239
Db
41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60

QY 240 GTGTGGAAGCCGGGAGACATGAGGTGGAATAACTCTCTGGAAGGGAGGCCGTGTGAGCGG 299
Db
61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80

QY 300 GTCCTGACCAGTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTTCGGGTGAACCTG 359
Db
81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100

QY 360 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db
101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120

QY 420 AGAAATGAGGCTGGTTTATCTCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG 479
Db
121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140

QY 480 GACAGTGACGGGGAATGGCACCGGCCAAAGCCATCATATAACGTCTTCCCTGATGGGAAA 539
Db
141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160

QY 540 CCTTTTCCTCACCAACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCTTCCACACTT 599
Db
161 PropheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180

QY 600 GGTCAGTATTTCCAGAAATTTGGACGATGTTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db
181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200

QY 660 GTGACACTTGGGCTCAACTCATGGAAGTGAAGTCTCTACAGAAGACATGGACGGGCATAT 719
Db
201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220

QY 720 GTTCCCATCGCAAGTGAAGATGTGTACGTGGTAACAGATCAGATTCTCTGTGTTTGTG 779
Db
221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240

QY 780 ACTATGTTCCAGAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAAGATCTCCCC 838
Db
241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260

QY 839 ATTATGTTTGATGTCCTGATTCATGATCTAGCCACTTCTCTCAATTATTCTACATTAAAC 898
Db
261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280

QY 899 TACAAGTGGAGTTCGGGGATAATACTGGCTGTTTGTTCACCAATCATACTACTGTGAAT 958
Db
281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300

QY 959 CACACGTATGTGCTCAATGGAACCTTTCAGCCTTAACCTCACTGTGAAAAGTCGACACCA 1018
Db
301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320

QY 1019 GGACCTTGTCCGCCACCGCCACCACCCAGACCTTCAAACCCACCCCTTCTTTAGGA 1078
Db
321 GlyProCysProProProProProProProProProProProProProProProProPro 340

QY 1079 CCTGCTGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACGCCAGATTAAAC 1138
Db
341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360

QY 1139 AGATATGGCCACTTTCAAGCCACCATCACAATTTAGAGGGAATCTTAGAGGTTAACATC 1198
Db
361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380

QY 1199 ATCCAGATGACAGACGTCCTGATGCGGTCGCATGGCCTGAAAGCTCCCTAATAGACTTT 1258
Db
381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400

QY 1259 GTCGTGACCTGCCAAGGGAGCATTCACCGAGGCTCTGTACCATCATTTCTGACCCACC 1318
Db
401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420

QY 1319 TCGGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCTG 1378
Db
421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440

QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCTTGGGGAT 1438

Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
Qy 1439 GACACAAAGCCTGGCTCTCACGAGCACCTGATTCTGTTCTGACAGAGACCCAGCCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
Qy 1499 CCTTTAAGGATGGCAAAACAGTCCCTGATCTCCGTTGGCTGCTGGCCATATTTGTCACT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
Qy 1559 GTGATCTCCCTCTTGGTGTACAAAAACACAAAGGAATACAAACCAATAGAAAATAGTCCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
Qy 1619 GGGAAATGTGTCAGAAAGCAAGGCCTGAGTGTCTTCTCAACCGTGCAGAAAGCCGTGTTT 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
Qy 1679 TTCCCGGGAACACAGGAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 3
ABB78200
ID ABB78200 standard; protein; 560 AA.
XX ABB78200;
XX
DT 25-NOV-2002 (first entry)
XX
DE Amino acid sequence of human HGFIN.
XX
KW Human; cell differentiation; white blood cell; bone marrow cell;
KW haematopoietic growth factor inducible neurokin-1; HGFIN;
KW progenitor proliferation; acute myeloid leukemia; non-Hodgkin's disease;
KW acute lymphocytic leukemia; chronic myeloid leukemia;
KW chronic lymphocytic leukemia; Hodgkin's disease.
XX

OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 1..485 /note= "extracellular domain"
FT Misc-difference 256 /note= "Leu encoded by CC"
FT Domain 485..508 /note= "transmembrane domain"
FT Domain 509..560 /note= "intracellular domain"
XX
PN WO200262947-A2.
XX
PD 15-AUG-2002.
XX
PF 20-OCT-2001; 2001WO-US050204.
XX
PR 20-OCT-2000; 2000US-0241881P.
XX
PA (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX
PI Pranela R;
XX
DR WPI; 2002-657531/70.
DR N-PSDB; ABQ78551.
XX
PT Hematopoietic growth factor inducible neurokin-1 type polypeptide and
PT polynucleotide for treating a disease associated with abnormal bone
PT marrow cell differentiation or proliferation, e.g. leukemia.
XX
PS Claim 12; Page 123-125; 125pp; English.
XX
CC The present sequence represents haematopoietic growth factor inducible

CC neurokin-1 type (HGFIN) polypeptide. HGFIN induces white blood cell
CC differentiation and inhibits progenitor proliferation. HGFIN polypeptides
CC and polynucleotides are useful for treating a disease associated with
CC abnormal bone marrow cell differentiation or proliferation, especially
CC acute myeloid leukemia, acute lymphocytic leukemia, chronic myeloid
CC leukemia, chronic lymphocytic leukemia, Hodgkin's and non-Hodgkin's
CC disease
XX
SQ Sequence 560 AA;
Alignment Scores: 4.82e-295 Length: 560
Pred. No.: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 5 Gaps: 0
US-10-039-272-1 (1-2661) x ABB78200 (1-560)
Qy 60 ATGGAATGTCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
Qy 120 GCCGCCAAACGATTTCATGATGTCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
Qy 180 CACAATCAATTAATGGCTGCTTCTCTGATGAAAATGACTGGAATGAAAAACTCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
Qy 240 GTGTGGAAGCGGGAGACATGAGGTGGAAAAACTCCTGGAAGGGAGCGGTGTCAGGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80
Qy 300 GTCTGACCAAGTCACTCACAGCCCTCGTGGCTCAAATATAACATTTGCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
Qy 360 ATATTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
Qy 420 AGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
Qy 480 GACAGTGACGGGAAAAATGGCACCGGCCAAAGCCATCATAACTCTTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
Qy 540 CCTTTTCCTCACCACCCCGGATGGAGAGATGGAATTTTCATCTACGTCTTCCACACACTT 599
Db 161 PropheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
Qy 600 GGTCAGTATTTCCAGAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
Qy 660 GTGACACTTGGGCTCAACTCATGGAAGTGAAGTGTGTACGAGTGTCTACAGAAGACATGGACGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
Qy 720 GTTCCCATCGCACAAAGTGAAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
Qy 780 ACTATGTTCCAGAAAGAACCGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
Qy 839 ATTATGTTTGATGTCCTGATTTCATGATCCTAGCCACTTCTCAATTATTCTACCATTAAC 898

Db	261	 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280
QY	899	TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATACTGTGAAT	958
Db	281	TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300
QY	959	CACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCA	1018
Db	301	HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro	320
QY	1019	GGACCTTGTCCGCCACCGCCACCACCCAGACCTTCAAAACCCACCCCTTCTTTAGGA	1078
Db	321	GlyProCysProProProProProProProProProProProProProProProProSerLeuGly	340
QY	1079	CCTGCTGGTGACAACCCCTGGAGCTGAGTAGGATTCCCTGATGAAAACTGCCAGATTAAC	1138
Db	341	ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn	360
QY	1139	AGATATGGCCACTTTCAAGCCACCATCACAAATTGTAGAGGAATCTTAGAGGTTAAACATC	1198
Db	361	ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle	380
QY	1199	ATCCAGATGACAGACGCTCCTGATCCCGGTGCCATGGCCCTGAAAGCTCCCTAATAGACTTT	1258
Db	381	IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe	400
QY	1259	GTCTGACCTGCCAAGGGAGCATTCCCACGGAGGTCTGTACCATCATTTCTGACCCACC	1318
Db	401	ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr	420
QY	1319	TGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCTG	1378
Db	421	CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu	440
QY	1379	ACTGTGAGACGAACCTTCAATGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGAT	1438
Db	441	ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp	460
QY	1439	GACACAAGCCTGGCTCTCACGAGCACCCCTGATTCTGTTCCTGACAGACCCAGCCTCG	1498
Db	461	AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer	480
QY	1499	CCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACT	1558
Db	481	ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr	500
QY	1559	GTGATCTCCCTCTTGGTGTAACAAAAACACAGGAATACAAACCCATAGAAAAATAGTCCT	1618
Db	501	ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro	520
QY	1619	GGGAATGGTGCAGAAAGCAAGCCCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTC	1678
Db	521	GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe	540
QY	1679	TTCCCGGGAACACAGGAAAGGATCCCGTACTCAAAACCAAGAAATTTAAAGGAGTTTCT	1738
Db	541	PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer	560
RESULT 4			
ABB74961			
ID ABB74961 standard; protein; 560 AA.			

XX	
PN	WO200200174-A2.
XX	
PD	03-JAN-2002.
XX	
PF	28-JUN-2001; 2001WO-US021065.
XX	
PR	28-JUN-2000; 2000US-00606421.
PR	02-AUG-2000; 2000US-00630940.
PR	21-AUG-2000; 2000US-00643597.
PR	15-SEP-2000; 2000US-00662786.
PR	09-OCT-2000; 2000US-00685696.
PR	12-DEC-2000; 2000US-00735705.
PR	07-MAY-2001; 2001US-00850716.
XX	
PA	(CORI-) CORIXA CORP.
XX	
PI	Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI	Mneill PD, Fanger N, Retter MW, Marnerakis M, Fanger GR;
PI	Vedvick TS, Carter D, Watanabe Y, Peckham DW;
XX	
DR	WPI; 2002-090513/12.
XX	
PT	Polynucleotides encoding lung tumor polypeptides, useful for treating
PT	lung cancer or stimulating an immune response.
XX	
PS	Disclosure; Page 281-283; 374pp; English.
XX	
CC	The present invention describes human lung tumour proteins. Human lung
CC	tumour proteins and polynucleotides have cytostatic and immunostimulant
CC	activities, and can be used in vaccine production. Compositions
CC	comprising the lung tumour proteins, polynucleotides, antibodies, fusion
CC	proteins, T cell populations, or antigen presenting cells that express
CC	the lung tumour proteins are useful for treating lung cancer or
CC	stimulating an immune response. ABL48959 to ABL49300 and ABB74946 to
CC	ABB75070 represent sequences used in the exemplification of the present
CC	invention
XX	
SQ	Sequence 560 AA;
Alignment Scores:	
Pred. No.:	4.82e-295 Length: 560
Score:	2999.00 Matches: 559
Percent Similarity:	99.82% Conservative: 0
Best Local Similarity:	99.82% Mismatches: 1
Query Match:	61.84% Indels: 1
DB:	5 Gaps: 0

QY	420	AGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG	479	QY	1499	CCTTTAAGGATGGCAAAACAGTGCCTGATCTCCGTTGGCTGGCCATATTGTCACT	1558
Db	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140	Db	481	ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr	500
QY	480	GACAGTGACGGGAAATGGCACCGGCCAAAGCCATCATACGCTTCCCTGATGGGAA	539	QY	1559	GTGATCTCCCTCTTGGTGTAACAAAAACACAAGGAATACACCCCAATAGAAAATAGTCCT	1618
Db	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProAspGlyLys	160	Db	501	ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro	520
QY	540	CCTTTCTCACCACCCCGGATGGAGAAGATGGAATTTCATCTACGCTTCCACACACTT	599	QY	1619	GGGAATGTGTCAGAAGCAAGGCCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTT	1678
Db	161	ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu	180	Db	521	GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe	540
QY	600	GGTCAGTATTCCAGAAATTGGGACGATGTTTCAGTGAGAGTTTTCTGTGAACACAGCCAAT	659	QY	1679	TTCCCGGAAACCCAGGAAAAGGATCCGCTACTCAAAAACCAAGAAATTAAGGAGTTTCT	1738
Db	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200	Db	541	PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer	560
QY	660	GTGACACTTGGGCCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGCATAT	719	RESULT 5			
Db	201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220	ABP61881			
QY	720	GTTCCCATCGCAAGTGAAGATGTGTACGTGGTAACAGATCAGATTCCTGTGTTGTG	779	ID	ABP61881	standard; protein; 560 AA.	
Db	221	ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal	240	XX	ABP61881;		
QY	780	ACTATGTTCCAGAGAACGATCGAAATTTCATCCGACGAAACCTTCCC-AAAGATCTCCCC	838	AC			
Db	241	ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro	260	XX			
QY	839	ATTATGTTTGATGCTCCTGATTTCATGATCCTAGCCACTTCCTCAATATTCTACCATTAAC	898	DT	07-OCT-2002	(first entry)	
Db	261	IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280	XX			
QY	899	TACAAGTGGAGCTTCGGGGATAATACTAGCCCTGTTGTTTCCACCAATCATACTGTGAAT	958	DE		Human lung cancer associated protein sequence SEQ ID NO:225.	
Db	281	TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300	KW		Human; lung cancer; lung tumour; cytostatic; gene therapy; vaccine.	
QY	959	CACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCA	1018	XX	WO200247534-A2.		
Db	301	HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro	320	XX	20-JUN-2002.		
QY	1019	GGACCTTGTCGCCACCGCCACCCACCCAGACCTTCAAAACCCACCCCTCTTTAGGA	1078	PF	30-NOV-2001; 2001WO-US047576.		
Db	321	GlyProCysProProProProProProProProProProProProProProProProSerLeuGly	340	XX	12-DEC-2000; 2000US-00735705.		
QY	1079	CCTGCTGGTGACAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACTGCCAGATTAAC	1138	PR	07-MAY-2001; 2001US-00850716.		
Db	341	ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn	360	PR	28-JUN-2001; 2001US-00897778.		
QY	1139	AGATATGGCCACTTCAAGCCACCATCACAATTGTAGAGGGAATCTTAGAGGTTAATCATC	1198	XX	(CORI-) CORIXA CORP.		
Db	361	ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle	380	DR	WPI; 2002-583465/62.		
QY	1199	ATCCAGATGACAGACGTCCTGATGCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTT	1258	XX			
Db	381	IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe	400	CC			
QY	1259	GTCGTGACCTGCCAAGGGAGCATTCACCGGAGGTCTGTACCATCTTTCTGACCCCCACC	1318	CC			
Db	401	ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr	420	CC			
QY	1319	TGCGAGATCACCCAGAACACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTGCTG	1378	CC			
Db	421	CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu	440	CC			
QY	1379	ACTGTGAGACGAACCTTCAATGGGTCTGGACGCTACTGTGTGAACCTCACCTGGGGAT	1438	CC			
Db	441	ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp	460	CC			
QY	1439	GACACAAGCCTGGCTCTCAGGACGACCCCTGATTCTGTCTCCTGACAGACCCAGCCTCG	1498	CC			
Db	461	AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer	480	XX			
				SQ		Sequence 560 AA;	

Alignment Scores:

Pred. No.:	4.82e-295	Length:	560
Score:	2999.00	Matches:	559
Percent Similarity:	99.82%	Conservative:	0
Best Local Similarity:	99.82%	Mismatches:	1
Query Match:	61.84%	Indels:	1
DB:	5	Gaps:	0

US-10-039-272-1 (1-2661) x ABP61881 (1-560)

QY	60	ATGGAATGTCTCTACTATTTCTGGGATTTCTGCTCCIGGCTGCAAGATTGCCACTTGTAT	111
Db	1		
		MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp	20
QY	120	GCCGCCAAAACGATTTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG	179
Db	21		
		AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40
QY	180	CACAATCAATTAATGGCTGCTTCTGATGAAAAATGACTGGAAATGAAAAAACTCTACCCA	239
Db	41		
		HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60
QY	240	GTGTGGAAGCGGGGAGACATGAGGTGGAAAAAACTCCTGGAAGGAGGCCGTGTGCAGGCG	299
Db	61		
		ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla	80
QY	300	GTCCTGACCAGTGACTCACCAGCCCTCGTGGGCTCAAAATAAATGCGGTGCGGTGAACCTG	359
Db	81		
		ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100
QY	360	ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC	419
Db	101		
		IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120
QY	420	AGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTCAACTGGACAGCATGGTCAGAG	479
Db	121		
		ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140
QY	480	GACAGTGACGGGAAAAATGGCACCGGCCAAAGCCATCATAAACGTCTTCCCTGATGGGAA	539
Db	141		
		AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProAspGlyLys	160
QY	540	CCTTTTCCACACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCTTCCACACACTT	599
Db	161		
		ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu	180
QY	600	GGTCAGTATTTCCAGAAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAA	659
Db	181		
		GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200
QY	660	GTGACACTTGGGCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATAT	719
Db	201		
		ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220
QY	720	GTTCCCATCGCACAAGTGAAGATGTGTACGTGTAACAGATCAGATTTCCTGTGTTGTG	779
Db	221		
		ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal	240
QY	780	ACTATGTTCCAGAAGAACGATCGAAATTTCATCCGACGAAACCTTCCC-AAAGATCTCCCC	838
Db	241		
		ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro	260
QY	839	ATTATGTTTGATGTCCTGATTTCATGATCCTAGCCACTTCCTCAATTATTCTACCAATAAC	898
Db	261		
		IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280
QY	899	TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTGTTTCCACCAATCATACTGTGAAT	958
Db	281		
		TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300
QY	959	CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGACCA	1018
Db	301		
		HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro	320

PR	18-APR-2001;	2001US-0284770P.
PR	10-MAY-2001;	2001US-0290492P.
PR	09-NOV-2001;	2001US-0339245P.
PR	13-NOV-2001;	2001US-0350666P.
PR	29-NOV-2001;	2001US-0334370P.
PR	12-APR-2002;	2002US-0372246P.
XX		
PA	(EOSB-) EOS BIOTECHNOLOGY INC	
XX		
PA	Aziz N, Murray R;	
PI		
XX		
DR	WPI; 2003-093161/08.	
DR	N-PSDB; ABX76321.	

The invention relates to a method for detecting a lung cancer-associated transcript in a cell from a patient, comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridises to a sequence that is at least 80 % identical to a gene that exhibits increased or decreased expression in lung cancer samples. Lung cancer-associated polynucleotides and polypeptides are used for identifying a compound that modulates a lung cancer-associated polypeptide, for inhibiting proliferation of a lung cancer-associated cell to treat lung cancer in a patient and for treating a mammal having lung cancer by administering a modulatory compound identified. The methods are useful for treating lung cancer, such as small cell lung cancer, non-small cell lung cancer or other benign or precancerous lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides and polypeptides are useful for diagnostic purposes and as targets for screening for therapeutic compounds that modulate lung cancer, such as antibodies. Sequences ABU56408-ABU56745 represent lung cancer-associated polypeptides of the invention

Alignment Scores:		
Pred. No.:	4.82e-295	Length: 560
Score:	2999.00	Matches: 559
Percent Similarity:	99.82%	Conservative: 0
Best Local Similarity:	99.82%	Mismatches: 1
Query Match:	61.84%	Indels: 1
DB:	6	Gaps: 0

US-10-039-272-1 (1-2661) x ABU56592 (1-560)

QY	60	ATGGAATGTCTCTACTATTCTCTGGGATTTCTGCTCTCTGGCTGCAAGATGCCACCTTGAT	119
Db	1	MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp	20
QY	120	GCCGCACAAACGATTTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG	179
Db	21	AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40
QY	180	CACAATCAATTAAATGGCTGTCCTCTGATGAAAAATGACTGGAATGAAAAACTCTACCCA	239
Db	41	HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60
QY	240	GTGTGGAGCGGGGAGACATGAGGTGGAAAAAATCTCTGGAAGGGAGGCCCGTGTGCAGGCG	299
Db	61	ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla	80
QY	300	GTCTGACCCAGTGACTCACCGCCCTCGTGGGCTCAAAATATAACATTTTGGGTGAACCTG	359
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100

QY	360	ATATTCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC	419
Db	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120
QY	420	AGAAATGAGCGTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG	479
Db	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140
QY	480	GACAGTGACGGGAAAAATGGCACCGGCCAAAAGCCATCATAAACGTCTTCCCTGATGGGAAA	539
Db	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProAspGlyLys	160
QY	540	CCTTTTCCTCACCAACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCCTCCACACACTT	599
Db	161	ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu	180
QY	600	GGTCAGTATTTCCAGAAAATTGGGACGATGTTTCAGTGAGAGTTTTCTGTGAACACAGCCAAT	659
Db	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200
QY	660	GTGACACTTGGGCCTCAACTCATGGAAGTCACTGTCTACAGAAGACATGGACGGGCATAT	719
Db	201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220
QY	720	GTTCCCATCGCACAAAGTGAAGATGIGTACGTGGTAAACAGATCAGATTCCTGTGTTTGTG	779
Db	221	ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal	240
QY	780	ACTATGTTCCAGAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC	838
Db	241	ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro	260
QY	839	ATTATGTTTGATGTCCTGATTTCATGATCCTAGCCACTTCCTCAATTATTCTACCATTAAC	898
Db	261	IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280
QY	899	TACAAGTGGAGCTCGGGGATAATACTGGCCTGTTTGTTCACCACTACTGTGAAT	958
Db	281	TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300
QY	959	CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGTCGACGACCA	1018
Db	301	HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaAlapro	320
QY	1019	GGACCTTGTCGGCCACCGCCACCACCCAGACCTTCAAAACCCACCCCTTCTTTAGGA	1078
Db	321	GlyProCysProProProProProProProProProProProProProProProProSerLysProThrProSerLeuGly	340
QY	1079	CCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACTGCCAGATTAAC	1138
Db	341	ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn	360
QY	1139	AGATATGGCCACTTCAAGCCACCACCATCACAAATTGTAGAGGGAATCTTAGAGGTTAACATC	1198
Db	361	ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle	380
QY	1199	ATCCAGATGACAGACGTCCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTT	1258
Db	381	IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe	400
QY	1259	GTCGTGACCTGCCAAGGAGACATTCACCACGGAGGTCGTACCATCATTTCTGACCCACC	1318
Db	401	ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr	420
QY	1319	TGGGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCTG	1378
Db	421	CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu	440
QY	1379	ACTGTGAGACGAACCTTCAATGGGTCTGGGACGTAAGCTGTGTGAACCTCACCCCTGGGGAT	1438
Db	441	ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp	460
QY	1439	GACACAAAGCCTGGCTCTCACGAGCACCCCTGATTTTCTGTCTCTGCAGAGACCCAGCCTCG	1498

||||| 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY CCTTTAAGGATGGCAACAGTGCCTTGATCTCCGTTGGTGCTTGGCCATATTTGTCACT 1558
Db ||||| 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTTGGTGTAACAAAAACACAAAGGAATACAAACCCCAATAGAAAAATAGTCCT 1618
Db ||||| 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGGAAATGTGTCAGAAAGCAAGGCCTGAGTGTCTTCTCAACCGTGC AAAAGCCGTGTTTC 1678
Db ||||| 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGGAACACAGGAAAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db ||||| 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 7
ABG72962
ID ABG72962 standard; protein; 560 AA.
XX
AC ABG72962;
XX
DT 08-APR-2003 (first entry)
XX
DE Human osteoactivin homologue, nmb.
XX
KW Human; osteoactivin; osteopathic; antiinflammatory; vaccine;
KW bone cell differentiation stumulator; gene therapy; bone formation;
KW osteoporosis; periodontal disease; ectopic bone formation; osteopetrosis;
KW bone disorder; osteogenesis; enzyme.
XX
OS Homo sapiens.
XX
PN US2002151486-A1.
XX
PD 17-OCT-2002.
XX
PF 30-AUG-2001; 2001US-00943075.
XX
PR 30-AUG-2000; 2000US-0229006P.
XX
PA (POPO/) POPOFF S N.
PA (SAFA/) SAFADI F F.
PA (OWEN/) OWEN T A.
PA (SMOC/) SMOCK S L.
XX
PI Popoff SN, Safadi FF, Owen TA, Smock SL;
XX
DR WPI; 2003-182528/18.
XX
PT Novel isolated osteoactivin protein and gene encoding the protein, useful
PT for stimulating bone differentiation and for treating bone disorders
PT including osteoporosis and periodontal disease.
XX
PS Claim 24; Fig 2B; 38pp; English.
XX
CC The invention describes an isolated and substantially pure osteoactivin
CC protein (I) that stimulates bone cell differentiation. A therapeutic
CC composition comprising (I), the polynucleotide (I) encoding (I) or a
CC biologically active fragment of (I) is useful for stimulating bone
CC formation in a mammal. The composition is also useful for treating a bone
CC disorder including osteoporosis and periodontal disease. A second
CC therapeutic composition comprising an anti-(I)-antibody or an agent that
CC inhibits osteoactivin-mediated bone formation is useful for inhibiting
CC formation in a mammal. The second composition is also useful for treating
CC ectopic bone formation and osteopetrosis. (I) and (II) are also useful
CC for developing novel therapeutic compositions for bone disorders, and for
CC stimulating osteogenesis. The polynucleotide is also useful in gene
CC therapy. This is the amino acid sequence of human nmb, a homologue of the
CC rat osteoactivin of the invention used in the creation of antibodies for

CC detection of nmb and osteoactivin
XX
SQ Sequence 560 AA;
Alignment Scores:
Pred. No.: 4.82e-295 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 6 Gaps: 0
US-10-039-272-1 (1-2661) x ABG72962 (1-560)
QY 60 ATGGAATGTCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db ||||| 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db ||||| 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAAATGGCTGGTCTTCTGATGAAAATGACTTGGAAATGAAAACTCTACCCA 239
Db ||||| 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAAAAACTCCTGGAAGGGAGGCGGTGTCAGGCG 299
Db ||||| 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80
QY 300 GTCTTGACCAAGTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTGCGGTGAACCTG 359
Db ||||| 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAAGAACTGC 419
Db ||||| 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG 479
Db ||||| 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGAAAAATGGCACCGGCCAAAAGCCATCATACGCTTCCCTGATGGGAAA 539
Db ||||| 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACCACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCTTCCACACACTT 599
Db ||||| 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAGTATTTCCAGAAATTGGGACCATGTTTCAGTGAGAGTTTCTGTGAAACACAGCCAAT 659
Db ||||| 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCCTCAACTCATGGAAGTGACTGTCTACAGAAACATGGACGGGCATAT 719
Db ||||| 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAAGATGTGTACGTGGTAACAGATCAGATTCTCTGTGTTGTG 779
Db ||||| 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAGAACCGATCGAAATTTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db ||||| 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTGATGTCTCTGATTTCATGATCCTAGCCACTTCTCCTCAATTATTCTACCATTAAC 898
Db ||||| 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCCTGTTTGTTCACCAATCATACTGTGAAT 958
Db ||||| 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300

QY 360 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db |||||
101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGCTGTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG 479
Db |||||
121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGAAAATGGCACCGGCCAAAGCCATCATAACGTCCTCCCTGATGGAAA 539
Db |||||
141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCCTCACACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCCTCCACACACTT 599
Db |||||
161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAGTATTTCCAGAAATTGGGACGATGTTCACTGAGAGTTTCTGTGAACACAGCCAAT 659
Db |||||
181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCCTCAACTCATGGAAGTGAAGTGTCTACAGAAGACATGGACGGGCATAT 719
Db |||||
201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTGTG 779
Db |||||
221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db |||||
241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTGATGTCTGATTCATGATCCTAGCCACTTCCTCAATTATTCTACCATTAAC 898
Db |||||
261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATACTGTGAAT 958
Db |||||
281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCA 1018
Db |||||
301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTGTGCCGCCACCGCCACCCAGCCAGACCTTCAAAACCCACCCCTTCTTTTAGGA 1078
Db |||||
321 GlyProCysProProProProProProProProProProProProProProProProProPro 340
QY 1079 CCTGTGTTGACAAACCCCTGGAGCTGAGTAGGATTCCTGTATGAAACTGCCAGATTAAAC 1138
Db |||||
341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTCAAGCCACCATCACAAATTGTAGAGGGAATCTTAGAGTTAAACATC 1198
Db |||||
361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCAGATGACAGACGTCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTT 1258
Db |||||
381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACCTGCCAAGGGAGCATTCACCAGGAGTCTGTACCATCATTTCTGACCCCAAC 1318
Db |||||
401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCTG 1378
Db |||||
421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGACGACGAACCTTCAATGGGTCTGGGACGTAAGTGTGAACCTCACCCCTGGGGAT 1438
Db |||||
441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460

QY 1439 GACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCCTGACAGAGACCCAGCCTCG 1498
Db |||||
461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGGATGGCAAAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACT 1558
Db |||||
481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTTGGTGTACAAAAACACAAAGGAATACACCCCAATAGAAAATAGTCCT 1618
Db |||||
501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGAATGTGTCAGAAAGCAAAGGCCCTGAGTGTCTTTCTCAACCCGTGCAAAAGCCGTGTTT 1678
Db |||||
521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGGAACCCAGGAAAAGGATCCGCTACTCAAAAAACCAAGATTTAAAGGAGTTTCT 1738
Db |||||
541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560
RESULT 9
ADA28315
ID ADA28315 standard; protein; 560 AA.
XX
AC ADA28315;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human lung tumour L528S protein.
XX
KW cancer; lung cancer; gene therapy; vaccine; human;
KW lung squamous cell carcinoma.
XX
OS Homo sapiens.
XX
PN US2003064947-A1.
XX
PD 03-APR-2003.
XX
PF 30-NOV-2001; 2001US-00007700.
XX
PR 18-MAR-1998; 98US-00040802.
PR 27-JUL-1998; 98US-00123912.
PR 22-DEC-1998; 98US-00221107.
PR 02-APR-1999; 99US-00285479.
PR 17-DEC-1999; 99US-00466396.
PR 30-DEC-1999; 99US-00476496.
PR 10-JAN-2000; 2000US-00480884.
PR 22-FEB-2000; 2000US-00510376.
PR 04-APR-2000; 2000US-00542615.
PR 28-JUN-2000; 2000US-00606421.
PR 02-AUG-2000; 2000US-00630940.
PR 21-AUG-2000; 2000US-00643597.
PR 15-SEP-2000; 2000US-00662786.
PR 09-OCT-2000; 2000US-00685696.
PR 12-DEC-2000; 2000US-00735705.
PR 07-MAY-2001; 2001US-00850716.
PR 28-JUN-2001; 2001US-00897778.
XX
PA (CORI-) CORIXA CORP.
XX
PI Wang T, Wang A, Skeiky YAW, Li SX, Kalos MD, Henderson RA;
PI Mcneill PD, Fanger N, Retter MW, Durham M, Fanger GR, Vedvick TS;
PI Carter D, Watanabe Y, Peckham DW, Cai F, Foy TW;
XX
DR WPI; 2003-540798/51.
XX
PT New isolated polynucleotides and polypeptides useful for diagnosing,
PT preventing and/or treating cancer, particularly lung cancer.
XX
PS Example 7; Page 194-195; 296pp; English.
XX
CC The invention describes isolated polynucleotides and polypeptides useful

CC for diagnosing, preventing and/or treating cancer, particularly lung
CC cancer. A new isolated polynucleotide comprises: any of the 22 fully
CC defined nucleotide sequences (e.g. 1012, 900 or 2773 bp) given in the
CC specification; complements of the nucleotide sequences cited above; at
CC least 10 contiguous residues of the nucleotide sequences cited above; a
CC sequence that hybridise to any of the nucleotide sequences under highly
CC stringent conditions; a sequence that is at least 75 or 90% identical to
CC the above nucleotide sequences; or degenerate variants of the above
CC nucleotide sequences. The composition and methods are useful in
CC diagnosing, preventing and/or treating cancer, particularly lung cancer,
CC in gene therapy and in vaccines. This is the amino acid sequence encoded
CC by a human lung tumour cDNA isolated from a lung squamous cell carcinoma
CC that may be useful in the diagnosis and treatment of lung cancer and
CC other disorders.
XX
SQ Sequence 560 AA;

Alignment Scores: 4.82e-295 Length: 560
Pred. No.: 2999.00 Matches: 559
Score: 99.82% Conservative: 0
Percent Similarity: 99.82% Mismatches: 1
Best Local Similarity: 61.84% Indels: 1
Query Match: 7 Gaps: 0
DB:

US-10-039-272-1 (1-2661) x ADA28315 (1-560)

QY 60 ATGGATGCTCTACTATTTCCTGGGATTTCTGCTCCTGGTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTCATGATGTCTGGGCAATGAAAGACCTTCGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAAATGGCTGCTTCTGATGAAATGACTGGAATGAAATAAATCTACCCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAATAAATCCTGGAAAGGAGCGCGTGTGAGGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80
QY 300 GTCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCTCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGCTGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGAAAATGGCACCGGCCCAAGCCATCATACGCTCTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACCAACCCCGGATGGGACGATGTTTCACTACGCTCTCCCTGCCACACTT 599
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAATTTCCAGAAAATGGGACGATGTTTCACTGAGAGTTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAGTGAAGATGTGTACTCGGTAAACAGATCAGATTCCCTGTTTGTG 779

Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAGAACGATCGAAAATTTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTGATGTCTGATTCATGATCCTAGCCACTTCTCCTCAATTATTCTACCAATTAA 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCCTGTTTGTTCACCAATCATACTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTGCTCAATGGAACCTTCAGCCCTTAACCTCAGCTGTGAAAGCTGCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCACCCACACCCAGACCTTCAAAAACCCACCTTCTTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProProProSerLeuGly 340
QY 1079 CCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACTGCCAGATTAA 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTCAAGCCACCATCACAAATGTAGAGGAATCTTAGAGGTTAACATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGCTCCTGATGCCGCTGCCATGGCCTGAAAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACCTGCCAAGGAGGACATTCACCGAGGCTGTGTACCATCATTTCTGACCCACC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TGCAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCTGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACACAGCCTGGCTCTCACGAGCACCTGATTTCTGTTCCTGACAGAGACCCAGCCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGGATGGCAACAGTGGCCCTGATCTCCGTTGGCTGCTTGGCCATATTGTCACT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTTGGTGTACAAAAACACAAGGAATACAAACCAATAGAAAATAGTCCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGGATGTGGTCAGAAAGCAAGGCTGAGTGTCTTTCTCAACCGTCGCAAAAGCCGTGTT 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGGAACACAGGAAAGGATCCGCTACTCAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 10
ADD78235
ID ADD78235 standard; protein; 572 AA.
XX
AC ADD78235;
XX
DT 29-JAN-2004 (first entry)

QY 1076 -----GGACCTGCTGGTGACAAACCCCTGGAG 1102
Db 341 ThrThrLeuLysSerTyrAspSerAsnThrProGlyProAlaGlyAspAsnProLeuGlu 360
QY 1103 CTGAGTAGGATTCTGATGAAACTGCCAGATTAACAGATATGGCCACTTTCAAGCCACC 1162
Db 361 LeuSerArgIleProAspGluAsnCysGlnIleAsnArgTyrGlyHisPheGlnAlaThr 380
QY 1163 ATCACAAATTGTAGAGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCTGTATG 1222
Db 381 IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet 400
QY 1223 CCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGCCAAGGGAGCATT 1282
Db 401 ProValProTrpProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle 420
QY 1283 CCCACGGAGGTCTGTACCATCATTTCTGACCCCCACCTGCGAGATCACCCAGAACACAGTC 1342
Db 421 ProThrGluValCysThrIleIleSerAspProThrCysGluIleThrGlnAsnThrVal 440
QY 1343 TGCAGCCCTGTGGATGTGGATGAGATGTCTGCTGACTGTGAGACGACCACTTCAATGGG 1402
Db 441 CysSerProValAspValAspGluMetCysLeuLeuThrValArgThrPheAsnGly 460
QY 1403 TCTGGGACGTACTGTGTGAACCTCACCTGGGGGATGACACAAGCTGGCTCTCACGAGC 1462
Db 461 SerGlyThrTyrCysValAsnLeuThrLeuGlyAspThrSerLeuAlaLeuThrSer 480
QY 1463 ACCCTGATTCTGTTCTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCC 1522
Db 481 ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla 500
QY 1523 CTGATCTCCGTGGCTGCTTGGCCATATTGTCACTGTGATCTCCCTCTTGGTGACAAA 1582
Db 501 LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuLeuValTyrLys 520
QY 1583 AAACACAAGGAATACAAACCCCAATAGAAATAGTCCTGGGAATGTGTCAGAAAGCAAGGC 1642
Db 521 LysHisLysGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerLysGly 540
QY 1643 CTGAGTGCTTCTTCAACCGTGCAAAAGCCGTGTTCTTCCCGGGAACACAGGAAAGGAT 1702
Db 541 LeuSerValPheLeuAsnArgAlaLysAlaValPhePheProGlyAsnGlnGluLysAsp 560
QY 1703 CCGCTACTCAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db 561 ProLeuLeuLysAsnGluGluPheLysGlyValSer 572

RESULT 11
AAU83612
ID AAU83612 standard; protein; 572 AA.

AC AAU83612;

DT 08-MAY-2002 (first entry)

DE Human PRO protein, Seq ID No 42.

KW Human; secreted protein; PRO; tumour; lung cancer; colon cancer;
KW breast cancer; prostate tumour; rectal tumour; liver tumour;
KW pericyte cell proliferation; chondrocyte cell proliferation;
KW tumour necrosis factor-alpha.

OS Homo sapiens.

PN WO200208288-A2.

XX 31-JAN-2002.

PF 29-JUN-2001; 2001WO-US021066.

PR 20-JUL-2000; 2000US-0219556P.

PR 25-JUL-2000; 2000US-0220585P.
PR 25-JUL-2000; 2000US-0220605P.
PR 25-JUL-2000; 2000US-0220607P.
PR 25-JUL-2000; 2000US-0220624P.
PR 25-JUL-2000; 2000US-0220638P.
PR 25-JUL-2000; 2000US-0220664P.
PR 25-JUL-2000; 2000US-0220666P.
PR 26-JUL-2000; 2000US-0220893P.
PR 28-JUL-2000; 2000WO-US020710.
PR 01-AUG-2000; 2000US-0222425P.
PR 22-AUG-2000; 2000US-0227133P.
PR 23-AUG-2000; 2000WO-US023522.
PR 24-AUG-2000; 2000WO-US023328.
PR 10-NOV-2000; 2000WO-US030873.
PR 28-NOV-2000; 2000US-0253646P.
PR 01-DEC-2000; 2000WO-US032678.
PR 20-DEC-2000; 2000US-00747259.
PR 20-DEC-2000; 2000WO-US034956.
PR 28-FEB-2001; 2001WO-US006520.
PR 01-MAR-2001; 2001WO-US006666.
PR 22-MAR-2001; 2001US-00816744.
PR 10-MAY-2001; 2001US-00854208.
PR 10-MAY-2001; 2001US-00854280.
PR 25-MAY-2001; 2001WO-US017092.
XX

(GETH) GENENTECH INC.

Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

WPI; 2002-172001/22.
N-PSDB; ABK33556.

One hundred and twenty two nucleic acids encoding PRO polypeptides,
useful for treating a PRO related disorder and for diagnosing tumors such
as lung cancer, colon cancer, breast tumor, prostate tumor, rectal tumor
or liver tumor.

Claim 11; Fig 42; 359pp; English.

The invention relates to one hundred and twenty two nucleic acids
encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides
encode human secreted proteins. The PRO nucleic acids, polypeptides,
agonists and antagonists are useful for treating a PRO related disorder.
The PRO polypeptides are useful for diagnosing tumours, especially lung
cancer, colon cancer, breast tumour, prostate tumour, rectal tumour or
liver tumour. The PRO polypeptides are useful for stimulating the
proliferation of, or gene expression, in pericyte cells, for stimulating
the proliferation or differentiation of chondrocyte cells, for
stimulating the release of tumour necrosis factor-alpha from human blood,
for stimulating or inhibiting the proliferation of normal human dermal
fibroblast cells. The PRO polypeptide may also be used as molecular
weight markers and for tissue typing. The PRO nucleic acids have
applications in molecular biology, including use as hybridisation probes,
and in chromosome and gene mapping. AAU83592-AAU83713 represent human PRO
protein sequences of the invention

SQ Sequence 572 AA;

Alignment Scores:

pred. No.: 5,26e-293 Length: 572
Score: 2979.00 Matches: 558
Percent Similarity: 97.55% Conservative: 0
Best Local Similarity: 97.55% Mismatches: 2
Query Match: 61.42% Indels: 13
DB: 5 Gaps: 1

US-10-039-272-1 (1-2661) x AAU83612 (1-572)

QY 60 ATGGAATGCTCTACTATTCTCTGGGATTTCGCTGGTGCAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20

QY 120 GCCGCCAAACGATTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAAATGGCTGGTCTTCTGATGATAAATGACTGGAATGAAAACTCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAAGCGGGGAGACATGAGGTGGAAAACTCCTGGAAAGGAGGCCGTGTGCAGGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80
QY 300 GTCCTGACCACTGACTCACCGCCCTCGTGGGCTCAAAATATAACATTTGCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGCTGGTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGGAANAATGGCACCGGCCAAACCCATCATAAACGTCTTCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCTACACCCCGGATGGAGAAGATGGAATTTCACTACGTCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAAGTATTTCCAGAAAATTTGGACGATGTTCAAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAAGATGTGTACGTGGTAACAGATCAGATTCCTGTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCGAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTGATGTCCTGATTCAATGATCCTAGCCACTTCCTCAATATATTTACCATTAA 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTGTTTCCACCAATCATACTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCACCACCCAGACCTTCAAAACCCACCCCTTCTTTA--- 1075
Db 321 GlyProCysProProProProProProProProProProProProProSerLysProthrProSerLeuAla 340
QY 1076 -----GGACCTGTGCTGACAAACCCCTGGAG 1102
Db 341 ThrThrLeuLysSerTyrAspSerAsnThrProGlyProThrGlyAspAsnProLeuGlu 360
QY 1103 CTGAGTAGGATTCCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACC 1162
Db 361 LeuSerArgIleProAspGluAsnCysGlnIleAsnArgTyrGlyHisPheGlnAlaThr 380
QY 1163 ATCACAATTGTAGAGGGAATCTTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATG 1222

Db 381 IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet 400
QY 1223 CCGGTGCCATGGCTGAAAGCTCCCTAATAGACTTTTGTGCTGACCTGCCAAGGAGCATT 1282
Db 401 ProValProTrpProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle 420
QY 1283 CCCACGGAGGTCTGTACCATCATTTCTGACCCCACTGCGGAGATCACCCAGAACACAGTC 1342
Db 421 ProThrGluValCysThrIleIleSerAspProThrCysGluIleThrGlnAsnThrVal 440
QY 1343 TGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTGCTGACTGTGAGACGAACCTTCAATGGG 1402
Db 441 CysSerProValAspValAspGluMetCysLeuLeuThrValArgArgThrPheAsnGly 460
QY 1403 TCTGGGACGTACTGTGTGAACCTCACCGCTGGGGGATGACACAAGCCCTGGCTCTCACGAGC 1462
Db 461 SerGlyThrTyrCysValAsnLeuThrLeuGlyAspThrSerLeuAlaLeuThrSer 480
QY 1463 ACCCTGATTTCTGTTCTCTGACAGAGACCCAGCCCTCGCCTTTAAGGATGCCAAACAGTGCC 1522
Db 481 ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla 500
QY 1523 CTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAA 1582
Db 501 LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuIleuValTyrLys 520
QY 1583 AAACACAAGGAATACAACCCAATAGAAAATAGTCTCTGGGAATGTGGTCAGAAAGCAAGGC 1642
Db 521 LysHisLysGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerLysGly 540
QY 1643 CTGAGTGTCTTCTCAACCCGTGCAAAAGCCGTGTCTTCCCGGGAACACAGGAAAAGGAT 1702
Db 541 LeuSerValPheLeuAsnArgAlaLysAlaValPhePheProGlyAsnGlnGluLysAsp 560
QY 1703 CCGCTACTCAAAAAACCAAGAATTTAAAGGAGTTTCT 1738
Db 561 ProLeuLeuLysAsnGlnGluPheLysGlyValSer 572
RESULT 12
ABU80759
ID ABU80759 standard; protein; 572 AA.
XX
AC ABU80759;
XX
DT 23-JUN-2003 (first entry)
XX
DE Human PRO polypeptide #21.
XX
KW Human; PRO polypeptide; secreted and transmembrane protein;
KW anti-PRO antibody; diagnostic assay; gene expression; tumour; cytostatic.
XX
OS Homo sapiens.
XX
PN US2003036635-A1.
XX
PD 20-FEB-2003.
XX
PF 28-AUG-2002; 2002US-00230163.
XX
PR 25-JUL-2000; 2000US-0220638P.
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
DR WPI; 2003-342045/32.
DR N-PSDB; ACA66861.
XX

PT One hundred and twenty two nucleic acids encoding PRO polypeptides,
PT useful for the manufacture of a medicament for diagnosing or treating
PT tumor.
XX
PS Claim 11; Fig 42; 314pp; English.
XX
CC The present invention relates to the isolation of novel human PRO
CC polypeptides, and the polynucleotide sequences encoding them. The PRO
CC polypeptides are secreted and transmembrane proteins. The PRO
CC polypeptides and polynucleotides are useful for preparing a medicament
CC useful in the diagnosis and treatment of tumours. Anti-PRO antibodies are
CC useful in diagnostic assays for PRO, by detecting its expression in
CC specific cells, tissues or serum, and for affinity purification of PRO
CC from recombinant cell culture or natural sources. ABU80739-ABU80860
CC represent the human PRO polypeptides of the invention. Note: The sequence
CC data for this patent was obtained in electronic format directly from the
CC USPTO web site at seqdata.uspto.gov/psipsDIDentry.html
XX
SQ Sequence 572 AA;

Alignment Scores:
Pred. No.: 5.26e-293 Length: 572
Score: 2979.00 Matches: 558
Percent Similarity: 97.55% Conservative: 0
Best Local Similarity: 97.55% Mismatches: 2
Query Match: 61.42% Indels: 13
DB: 6 Gaps: 1

US-10-039-272-1 (1-2661) x ABU80759 (1-572)

QY	60	ATGGAATGTC	TACTATTTCCTGGGATTTCTGCTCCTGGCTGC	AAGATTGCCACTTGAT	119
Db	1	MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuLeuAlaAlaArgLeuProLeuAsp	20		
QY	120	GCCGCCAACG	ATTTCATGATGTCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAG	179	
Db	21	AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40		
QY	180	CACAATCAAT	TAATGGCTGCTTCTGATGAAAAATGACTGGAATGAAAAA	CTCTACCCCA	239
Db	41	HisAsnGlnLeuAsnGlyTyrPheLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	60		
QY	240	GTGTGGAAGCGGGAGACATGAGGTGGAAAAA	CTCCTGGAAGGAGCGGTGTCAGGCG	299	
Db	61	ValTyrLysArgGlyAspMetArgTyrLysAsnSerTyrLysGlyArgValGlnAla	80		
QY	300	GTCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTTCCGGTGAACCTG	359		
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100		
QY	360	ATATTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC	419		
Db	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120		
QY	420	AGAAATGAGGTGGTTTATCTGCTGATCCATATGTTTACAAC	TGGACAGCATGGTCAGAG	479	
Db	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTyrPheAlaTyrPheSerGlu	140		
QY	480	GACAGTGACGGGAAAAATGGCACCGGCCAAAGCCATCATACGCTTCCCTGATGGGAAA	539		
Db	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys	160		
QY	540	CCTTTTCCTCACCCCGGATGGAGAGATGGAATTTTCATCTACGCTTCCACACACTT	599		
Db	161	ProPheProHisHisProGlyTyrArgArgTyrPhePheIleTyrValPheHisThrLeu	180		
QY	600	GGTCAGTATTTCCAGAAATTTGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAA	T	659	
Db	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200		
QY	660	GTGACACTTGGGCCTCAACTCATGGAAGTGAAGTGTCTACAGAAGACATGGACGGGCATAT	719		
Db	201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220		

QY	720	GTTCCTCATCGCACAAAGTGAAGATGTGTACGTGGTAACAGATCAGATTCTCTGTGTTGTG	779
Db	221	ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal	240
QY	780	ACTATGTTCCAGAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC	838
Db	241	ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro	260
QY	839	ATTATGTTTGATGTCCTGATTCATGATCCTAGCCACTTCCATCAATTATTCTACCATTAAC	898
Db	261	IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280
QY	899	TACAAGTGGAGCTTCGGGGATAACTGGCCTGTTTGTTCACCAATCATACTGTGAAT	958
Db	281	TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300
QY	959	CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCA	1018
Db	301	HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro	320
QY	1019	GGACCTTGTCCGCCACCGCCACCAACCCAGACCTTCAAAACCCACCCCTTCTTTA---	1075
Db	321	GlyProCysProProProProProProProProProProProProProProProProPro	340
QY	1076	-----GGACCTGCTGCTGACAAACCCCTGGAG	1102
Db	341	ThrThrLeuLysSerTyrAspSerAsnThrProGlyProThrGlyAspAsnProLeuGlu	360
QY	1103	CTGAGTAGGATTCTCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACC	1162
Db	361	LeuSerArgIleProAspGluAsnCysGlnIleAsnArgTyrGlyHisPheGlnAlaThr	380
QY	1163	ATCACAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATG	1222
Db	381	IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet	400
QY	1223	CCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTTGTGCTGACTGCCAAGGAGCATT	1282
Db	401	ProValProTyrProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle	420
QY	1283	CCCACGAGGTCTGTACCATCATTTCTGACCCCCACCTCGGAGATCACCCAGAACACAGTC	1342
Db	421	ProThrGluValCysThrIleIleSerAspProThrCysGluIleThrGlnAsnThrVal	440
QY	1343	TGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTCTGCTGACTGTGAGACCACTTCAATGGG	1402
Db	441	CysSerProValAspValAspGluMetCysLeuLeuThrValArgArgThrPheAsnGly	460
QY	1403	TCTGGGACGTACTGTGTGAACCTCACCTGGGGGATGACACAGCCTGGCTCTCACGAGC	1462
Db	461	SerGlyThrTyrCysValAsnLeuThrLeuGlyAspAspThrSerLeuAlaLeuThrSer	480
QY	1463	ACCCTGATTTCTGTTCTCTGACAGAGACCCAGCCCTCGCCTTTAAGGATGGCAAAACAGTGCC	1522
Db	481	ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla	500
QY	1523	CTGATCTCCGTTGGCTGCTTGGCCATATTGTCACTGTGATCTCCCTCTTGGTGACAAA	1582
Db	501	LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuLeuValTyrLys	520
QY	1583	AAACACAAGGAATACACCCCAATAGAAAATAGTCTCTGGGAATGTGTCAGAACAAAGGC	1642
Db	521	LysHisLysGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerLysGly	540
QY	1643	CTGAGTGTCTTCTCAACCGTGC AAAAGCCGTGTTCTTCCCGGGGAAACCCAGGAAAGGAT	1702
Db	541	LeuSerValPheLeuAsnArgAlaLysAlaValPhePheProGlyAsnGlnGluLysAsp	560
QY	1703	CCGCTACTCAAAAACCAAGAAATTTAAAGAGTTTCT	1738
Db	561	ProLeuLeuLysAsnGlnGluPheLysGlyValSer	572

RESULT 13
ABO33725
ID ABO33725 standard; protein; 572 AA.
XX AC ABO33725;
XX DT 17-SEP-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO9925.
XX
KW Human; secreted and transmembrane protein; PRO; cytostatic;
KW antiarthritic; osteopathic; gene therapy; TNF-Agonist-Alpha;
KW chondrocyte stimulator; pericyte stimulator; fibroblast modulator;
KW pharmaceutical; diagnostic; biosensor; bioreactor; tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW liver tumour; bone disorder; cartilage disorder; sports injury;
KW arthritis; wound.
XX
OS Homo sapiens.
XX
PN US2003045687-A1.
XX
PD 06-MAR-2003.
XX
XX 12-AUG-2002; 2002US-00218631.
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
DR WPI; 2003-512315/48.
DR N-PSDB; ACD68613.
XX
PT New genes, and its encoded secreted and transmembrane polypeptides,
PT useful for stimulating Tumor Necrosis Factor alpha, or chondrocyte or
PT pericyte proliferation, especially for treating lung tumors, arthritis or
PT wounds in a mammal.
XX
PS Claim 11; Fig 42; 314pp; English.
XX
CC The invention describes an isolated nucleic acid molecule comprising a
CC sequence with at least 80% identity to: (a) a nucleotide encoding any of
CC 122 PRO (secreted and transmembrane) polypeptides whose sequences are
CC fully defined in the specification; or (b) any of 122 nucleotide
CC sequences having e.g. 4834, 2504 or 1759 bp fully defined in the
CC specification; or the full length coding sequence of any these 122
CC nucleotide sequences. The PRO polypeptides or polynucleotides are useful
CC as pharmaceuticals, diagnostics, biosensors or bioreactors. These are
CC particularly useful for detecting tumours (e.g. lung tumour, colon
CC tumour, breast tumour, prostate tumour, rectal tumour, or liver tumour)
CC in a mammal, for stimulating the release of TNF-alpha from human blood,
CC for stimulating the proliferation or differentiation of chondrocyte
CC cells, for stimulating proliferation of pericyte cells, or for modulating
CC normal human dermal fibroblast proliferation. The PRO nucleic acid or
CC polypeptide is also useful for treating tumours or various bone and/or
CC cartilage disorders (e.g. sports injuries or arthritis), or wounds. The
CC PRO polypeptides are useful in drug screening, particularly as targets
CC for therapeutic intervention in these diseases, and in the diagnostic
CC determination of the presence of these diseases. The PRO polypeptides are
CC also useful as molecular weight markers, or for chromosome
CC identification. The PRO genes are useful as hybridisation probes, or for
CC screening libraries of human cDNA, genomic DNA or mRNA. The PRO genes may
CC also be used in gene therapy, particularly for replacing a defective
CC gene. This is the amino acid sequence of a novel human secreted and
CC transmembrane PRO polypeptide
XX
SQ Sequence 572 AA;

Alignment Scores:			
Pred. No.:	5.26e-293	Length:	572
Score:	2979.00	Matches:	558
Percent Similarity:	97.55%	Conservative:	0
Best Local Similarity:	97.55%	Mismatches:	2
Query Match:	61.42%	Indels:	13
DB:	6	Gaps:	1
US-10-039-272-1 (1-2661) x ABO33725 (1-572)			
QY	60	ATGGAATGCTCTACTATTCTCTGGGATTTCTGCTCTGGTGCAAGATTGCCACTTGAT	119
Db	1	MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp	20
QY	120	GCCGCCAAACGATTTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGAG	179
Db	21	AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40
QY	180	CACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGATGAAAACTCTACCCA	239
Db	41	HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60
QY	240	GTGTGGAGCGGGGAGACATGAGGTGGAAAAAATCCTGGAGGGAGCCGTGTGCAGGCG	299
Db	61	ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla	80
QY	300	GTCTGACCACTGACTCACCAAGCCCTCGTGGGCTCAAATATAACATTTCCGGTGAACCTG	359
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100
QY	360	ATATTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC	419
Db	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120
QY	420	AGAAATGAGGCTGGTTTATCTGCTGATCCCATATGTTTACAACGGACAGCATGGTCAGAG	479
Db	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140
QY	480	GACAGTGACGGGAAAATGGCACCGGCCAAAGCCATCATAACTCTTCCCTGATGGAAA	539
Db	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys	160
QY	540	CCTTTTCCTCACCAACCCCGGATGGAGAAGATGGAATTTTCATCTACGTTCTCCACACTT	599
Db	161	ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu	180
QY	600	GGTCAGTATTTCCAGAAATTTGGGACGATGTTTCAGTGAGAGATTTCTGTGAACACAGCCAAT	659
Db	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200
QY	660	GTGACACTTGGGCCTCAACTCATGGAAGTGACTGTCTACAGHACACATGGACGGGCATAT	719
Db	201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220
QY	720	GTTCCCATCGCACAACTGAAAGATGTGTACGTGGTTAACAGATCAGATTCCCTGTTTGTG	779
Db	221	ValProlleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal	240
QY	780	ACTATGTTCCAGAAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC	838
Db	241	ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro	260
QY	839	ATTATGTTTGATGTCTGATTCATGATCCTAGCCACTTCTCCTCAATTATTCTACCATTAAC	898
Db	261	IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280
QY	899	TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTGTTTCCACCAATCATACTGTGAAT	958
Db	281	TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300
QY	959	CACACGTATGTGCTCAATGGAACTTTCAGCCTTAACCTTCACTGTGAAAGCTGCAGCACCA	1018
Db	301	HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro	320

Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGAAAAATGGCACCGGCCAAAGCCATCATAAAGTCTTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCCTCACCAACCCCGGATGGAGAGATGGAATTTTCATCTACGTCTTCCACACACTT 599
Db 161 PropheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAAGTATTTCCAGAAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAAGATGTGTACGTGGTAAACAGATCAGATTCTCTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAGAACGATCGAAATTTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTGATGTCCTGATTTCATGATCCTAGCCACTTCCTCAATTATTTACCATTAAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCGCTGTTGTTTCCACCACATCATCTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCAACCCAGACCTTCAAAACCCACCTTCTTTA --- 1075
Db 321 GlyProCysProProProProProProProProProProProProSerLysProthrProSerLeuAla 340
QY 1076 -----GGACCTGTCTGGTGACAAACCCCTGGAG 1102
Db 341 ThrThrLeuLysSerTyrAspSerAsnThrProGlyProThrGlyAspAsnProLeuGlu 360
QY 1103 CTGAGTAGGATTCCTGATGAAAACCTGCCAGATTAAACAGATATGGCCACTTCAAGCCACC 1162
Db 361 LeuSerArgIleProAspGluAsnCysGlnIleAsnArgTyrGlyHisPheGlnAlaThr 380
QY 1163 ATCACAATTGTAGAGGGAATCTTAGAGGTTAATCATCATCCAGATGACAGACGTCCTGATG 1222
Db 381 IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet 400
QY 1223 CCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGCCAAGGGAGCATT 1282
Db 401 ProValProTrpProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle 420
QY 1283 CCCACGGAGGTCTGTACCATCATTTCTGACCCACCTGCGAGATCACCCAGAACACAGTC 1342
Db 421 ProThrGluValCysThrIleIleSerAspProThrCysGluIleThrGlnAsnThrVal 440
QY 1343 TGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGAACCTTCAATGGG 1402
Db 441 CysSerProValAspValAspGluMetCysLeuLeuThrValValArgArgThrPheAsnGly 460
QY 1403 TCTGGGACGTACTGTGTGAACCTCACCCCTGGGGGATGACACAAGCCTGGCTCTCACGAGC 1462
Db 461 SerGlyThrTyrCysValAsnLeuThrLeuGlyAspThrSerLeuAlaLeuThrSer 480

QY 1463 ACCCTGATTTCTGTTCTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGCAGAAACAGTGCC 1522
Db 481 ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla 500
QY 1523 CTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGACAAA 1582
Db 501 LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuLeuValTyrLys 520
QY 1583 AACACAAAGGAATACAAACCCCAATAGAAAATAGTCTCTGGGAATGTGGTCAGAAAGCAAGGC 1642
Db 521 LysHisLysGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerLysGly 540
QY 1643 CTGAGTGTCTTCTCAACCGTGCAGAAAGCCGTGTTCTTCCCGGAAACCCAGGAAAAGGAT 1702
Db 541 LeuSerValPheLeuAsnArgAlaLysAlaValPhePheProGlyAsnGlnGluLysAsp 560
QY 1703 CCGCTACTCAAAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db 561 ProLeuLeuLysAsnGlnGluPheLysGlyValSer 572
RESULT 15
ABJ72248
ID ABJ72248 standard; protein; 572 AA.
XX
AC ABJ72248;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human PRO9925 protein.
XX
KW PRO; proliferation; pericyte cell; TNF-alpha; blood; chondrocyte;
KW differentiation; dermal fibroblast; tumour; gene therapy; cytostatic.
XX
OS Homo sapiens.
XX
PN US2003050448-A1.
XX
PD 13-MAR-2003.
XX
PF 28-AUG-2002; 2002US-00230414.
XX
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
DR WPI; 2003-521818/49.
DR N-PSDB; ABT44246.
XX
PT New nucleic acid encoding for a PRO protein, useful for the manufacture
PT of a medicament for diagnosing or treating tumors or for measuring or
PT detecting expression of an associated gene.
XX
PS Claim 11; Fig 42; 315pp; English.
XX
CC The invention relates to a novel isolated nucleic acid encoding a fully
CC defined PRO polypeptide. The molecules of the invention may be useful for
CC stimulating proliferation or gene expression in pericyte cells or the
CC release of TNF-alpha from human blood. Other possible uses include the
CC stimulation or inhibition of chondrocyte proliferation or
CC differentiation, the stimulation of human dermal fibroblast cell
CC proliferation and the detection of the presence of a tumour within a
CC mammal. Furthermore, the nucleic acid may be useful for the manufacture
CC of a medicament for diagnosing or treating a tumour within a mammal or
CC for measuring or detecting the expression of an associated gene, as well
CC as during gene therapy. The current sequence is that of the human PRO
XX protein of the invention

SQ	Sequence 572 AA;				
Alignment Scores:					
Pred. No.:	5.26e-293	Length:	572		
Score:	2979.00	Matches:	558		
Percent Similarity:	97.55%	Conservative:	0		
Best Local Similarity:	97.55%	Mismatches:	2		
Query Match:	61.42%	Indels:	13		
DB:	6	Gaps:	1		
US-10-039-272-1 (1-2661) x ABJ72248 (1-572)					
QY	60	ATGGAATGTCTCTACTATTTCTCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT	119		
Db	1	MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuLeuAlaAArgLeuProLeuAsp	20		
QY	120	GCCGCCAAACGATTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG	179		
Db	21	AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40		
QY	180	CACAAATCAATTAAATGGCTGGTCTTCTGATGAAAAATGACTGGAATGAAAAACTCTACCCA	239		
Db	41	HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60		
QY	240	GTGTGGAAGCGGGAGACATGAGGTGGAAAAAATCTCTGGAAGGAGGCCGTGTGCAGGCG	299		
Db	61	ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla	80		
QY	300	GTCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTG	359		
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100		
QY	360	ATATTCCTTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC	419		
Db	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120		
QY	420	AGAAATGAGCGTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG	479		
Db	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140		
QY	480	GACAGTGACGGGAAAAATGGCACCGGCCAAAGCCATCATAAACGTCTTCCCTGATGGGAAA	539		
Db	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys	160		
QY	540	CCTTTTCTCACCAACCCCGGATGGAGAAAGATGGAATTTCACTACGTCTTCCACACACTT	599		
Db	161	ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu	180		
QY	600	GGTCAGTATTCCAGAAATTGGGACGATGTTCACTGAGAGTTTCTGTGAACACACAGCCAAT	659		
Db	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200		
QY	660	GTGACACTTGGGCCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATAT	719		
Db	201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220		
QY	720	GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCTCTGTGTTGTG	779		
Db	221	ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal	240		
QY	780	ACTATGTTCCAGAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC	838		
Db	241	ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro	260		
QY	839	ATTATGTTTGTGATCTCTGATTTCATGATCCTAGCCACTTCTCTCAATTATTCTACCATTAAC	898		
Db	261	IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280		
QY	899	TACAAGTGGAGCTTCGGGGATAATACTGGCCCTGTTGTTTCCACCAATCATACTCTGAAT	958		
Db	281	TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300		
QY	959	CACACGTATGTGCTCAATGGAACCTTCAGCCCTTAACCTCACTGTGAAAGCTGCAGCACCA	1018		

DB	301	HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro	320		
QY	1019	GGACCTTGTCCGCCACCCACACCACCCAGACCTTCAAAACCCACCCCTTCTTTTA---	1075		
DB	321	GlyProCysProProProProProProProProProProProProProProProProProPro	340		
QY	1076	-----GGACCTGCTGGTGACAACCCCTGGAG	1102		
DB	341	ThrThrLeuLysSerTyrAspSerAsnThrProGlyProThrGlyAspAsnProLeuGlu	360		
QY	1103	CTGAGTAGGATTCTGTATGAAAACTGCCAGATTAAACAGATTATGGCCACTTTCAAGCCACC	1162		
DB	361	LeuSerArgIleProAspGluAsnCysGlnIleAsnArgTyrGlyHisPheGlnAlaThr	380		
QY	1163	ATCACAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCTCTGATG	1222		
DB	381	IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet	400		
QY	1223	CCGCTGCCATGGCCTGAAAGCTCCCTAATAGACTTTTGTCTGACCTGCCAAGGGAGCAAT	1282		
DB	401	ProValProTrpProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle	420		
QY	1283	CCACCGGAGGTCTGTACCATCATTTCTGACCCACCTCGGAGATCACCCAGAACACAGTC	1342		
DB	421	ProThrGluValCysThrIleIleSerAspProThrCysGluIleThrGlnAsnThrVal	440		
QY	1343	TGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTGACTGTGAGACGAACTTCAATGGG	1402		
DB	441	CysSerProValAspValAspGluMetCysLeuLeuThrValArgArgThrPheAsnGly	460		
QY	1403	TCTGGGACGTACTGTGTGAACCTCACCCCTGGGGATGACAAAGCCTGGCTCTCACGAGC	1462		
DB	461	SerGlyThrTyrCysValAsnLeuThrLeuGlyAspAspThrSerLeuAlaLeuThrSer	480		
QY	1463	ACCTGATTTCTGTTCTCTGACAGACCCAGCCTCGCCTTTAAAGGATGGCAACAGTGCC	1522		
DB	481	ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla	500		
QY	1523	CTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGTATCTCCCTCTTGGTGTACAAA	1582		
DB	501	LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuLeuValTyrLys	520		
QY	1583	AAACACAAGGAATACAAACCCATAGAAAATAGTCTCTGGGAATGTGGTCAGAAGCAAAGGC	1642		
DB	521	LysHisLysGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerLysGly	540		
QY	1643	CTGAGTCTTTTCTCAACCGGTGCAAAAGCCGTGTTCTTCCCGGGAACCCAGGAAAAGGAT	1702		
DB	541	LeuSerValPheLeuAsnArgAlaLysAlaValPhePheProGlyAsnGlnGluLysAsp	560		
QY	1703	CCGCTACTCAAAAAACCAAGAATTTAAAGGAGTTTCT	1738		
DB	561	ProLeuLeuLysAsnGlnGluPheLysGlyValSer	572		

Search completed: September 10, 2004, 15:39:40
Job time : 236 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 9, 2004, 20:21:02 ; Search time 1483 Seconds
(without alignments)
7622.694 Million cell updates/sec

Title: US-10-039-272-1
Perfect score: 2661
Sequence: 1 cggcacgagggcccgagga.....aaaaaaaaaaaaaaaa 2661

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002s:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query %			ID	Description
	Score	Match	Length		
1	2661	100.0	2661	ABQ78551	Abq78551 Nucleotid
2	2547.8	95.7	2669	AAT69328	Aat69328 Murine me
3	2547.8	95.7	2669	AAT69318	Aat69318 Murine me
4	2547.8	95.7	2669	AAC55715	Aac55715 Human NMB
5	2547.8	95.7	2669	ABQ88185	Abq88185 Human ost
6	2547.8	95.7	2669	ABX76321	Abx76321 Lung canc
7	2505.6	94.2	2728	ADD78274	Add78274 Human CGD
8	2485	93.4	2683	ABK33556	Abk33556 cDNA enco
9	2485	93.4	2683	ACA66861	Aca66861 cDNA enco
10	2485	93.4	2683	ACD68613	Acd68613 Novel hum
11	2485	93.4	2683	ACA68517	Aca68517 Novel hum
12	2485	93.4	2683	ABT44246	Abt44246 Human PRO
13	2485	93.4	2683	ABT44529	Abt44529 Human PRO
14	2485	93.4	2683	ACD82196	Acd82196 Human sec
15	2485	93.4	2683	ABT43902	Abt43902 Human mem
16	2485	93.4	2683	ADB83531	Adb83531 Novel hum
17	2485	93.4	2683	ADB80637	Adb80637 Novel hum
18	2485	93.4	2683	ADB73178	Adb73178 Novel hum
19	2485	93.4	2683	ADB78260	Adb78260 Novel hum
20	2485	93.4	2683	ADB84908	Adb84908 Human PRO
21	2485	93.4	2683	ADB78014	Adb78014 Novel hum
22	2485	93.4	2683	ADB87080	Adb87080 Human PRO
23	2485	93.4	2683	ADB84662	Adb84662 Human PRO

24	2485	93.4	2683	9	ADB83777	Novel hum
25	2485	93.4	2683	9	ADB72932	Novel hum
26	2485	93.4	2683	9	ADC36770	Human PRO
27	2485	93.4	2683	9	ADC21760	Human PRO
28	2485	93.4	2683	9	ADC49791	Novel hum
29	2485	93.4	2683	9	ADC48990	Novel hum
30	2485	93.4	2683	9	ADC49507	Novel hum
31	2485	93.4	2683	9	ADC47368	Novel hum
32	2485	93.4	2683	9	ADC47113	Novel hum
33	2485	93.4	2683	9	ADC77988	Novel hum
34	2485	93.4	2683	9	ADD06223	Novel hum
35	2485	93.4	2683	9	ADC77742	Novel hum
36	2485	93.4	2683	9	ADD50705	Novel hum
37	2485	93.4	2683	9	ADD50951	Novel hum
38	2485	93.4	2683	9	ADD50432	Human PRO
39	2485	93.4	2683	9	ADD50186	Human PRO
40	2485	93.4	2683	9	ADD51197	Novel hum
41	2485	93.4	2683	10	ADC48744	Novel hum
42	2485	93.4	2683	10	ADE20915	Novel hum
43	2485	93.4	2683	10	ADE05759	Human PRO
44	2485	93.4	2683	10	ADD74988	Human PRO
45	2485	93.4	2683	10	ADD75734	Novel hum

ALIGNMENTS

RESULT 1
ABQ78551
ID ABQ78551 standard; DNA; 2661 BP.
XX
AC ABQ78551;
XX
DT 25-NOV-2002 (first entry)
XX
DE Nucleotide sequence of human HGFIN.
XX
KW Human; cell differentiation; white blood cell; bone marrow cell;
KW haematopoietic growth factor inducible neurokin-1; HGFIN;
KW progenitor proliferation; acute myeloid leukemia; non-Hodgkin's disease;
KW acute lymphocytic leukemia; chronic myeloid leukemia;
KW chronic lymphocytic leukemia; Hodgkin's disease; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 60..1741
FT /*tag= a
FT /product= "HGFIN"
FT /transl_except= (825..826,aa:Leu)
XX
XX
PN WO200262947-A2.
XX
XX
PD 15-AUG-2002.
XX
PF 20-OCT-2001; 2001WO-US050204.
XX
PR 20-OCT-2000; 2000US-0241881P.
XX
PA (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX
PI Pranela R;
XX
DR WPI; 2002-657531/70.
XX P-PSDB; ABB78200.
PT Hematopoietic growth factor inducible neurokin-1 type polypeptide and
PT polynucleotide for treating a disease associated with abnormal bone
PT marrow cell differentiation or proliferation, e.g. leukemia.
XX
PS Claim 2; Page 121-123; 125pp; English.
XX
CC The present sequence encodes human haematopoietic growth factor inducible

CC neurokin-1 type (HGFN) polypeptide. HGFN induces white blood cell
CC differentiation and inhibits progenitor proliferation. HGFN polypeptides
CC and polynucleotides are useful for treating a disease associated with
CC abnormal bone marrow cell differentiation or proliferation, especially
CC acute myeloid leukemia, acute lymphocytic leukemia, chronic myeloid
CC leukemia, chronic lymphocytic leukemia, Hodgkin's and non-Hodgkin's
CC disease
XX
SQ Sequence 2661 BP; 772 A; 586 C; 587 G; 716 T; 0 U; 0 Other;

Query Match 100.0%; Score 2661; DB 6; Length 2661;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCACGAGGCCCCAGAGGAATAAGTTAAACCTTGGTGCCTGCGTCGGTGAAGATTTCAGCA 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1 CGGCACGAGGCCCCAGAGGAATAAGTTAAACCTTGGTGCCTGCGTCGGTGAAGATTTCAGCA 60

QY 61 TGGAAATGCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGATG 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
61 TGGAAATGCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGATG 120

QY 121 CCGCCAAACGATTTTCATGATGTGCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAGC 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
121 CCGCCAAACGATTTTCATGATGTGCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAGC 180

QY 181 ACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTTACCCAG 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
181 ACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTTACCCAG 240

QY 241 TGTGGAAGCGGGAGACATGAGGTGGAAAAAATCCTCTGGAAGGAGGCCGTGTGCAGGCGG 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
241 TGTGGAAGCGGGAGACATGAGGTGGAAAAAATCCTCTGGAAGGAGGCCGTGTGCAGGCGG 300

QY 301 TCCTGACCCAGTGACTCACCAGCCCTCGTGGCTCAAATATAACATTTGCGGTGAACCTGA 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
301 TCCTGACCCAGTGACTCACCAGCCCTCGTGGCTCAAATATAACATTTGCGGTGAACCTGA 360

QY 361 TATTCCTTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGCA 420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
361 TATTCCTTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGCA 420

QY 421 GAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAGG 480
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
421 GAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAGG 480

QY 481 ACAGTGACGGGGAATAATGGCACCGGCCAAAGCCATCATAACTCTTCCCTGATGGGAAC 540
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
481 ACAGTGACGGGGAATAATGGCACCGGCCAAAGCCATCATAACTCTTCCCTGATGGGAAC 540

QY 541 CTTTTCCTCACCACCCCGGATGGAGAAGATGGAATTCATCTACGTCTTCCACACACTTG 600
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
541 CTTTTCCTCACCACCCCGGATGGAGAAGATGGAATTCATCTACGTCTTCCACACACTTG 600

QY 601 GTCAGTATTTCCAGAAATGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAATG 660
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
601 GTCAGTATTTCCAGAAATGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAATG 660

QY 661 TGACACTTGGGCCCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATATG 720
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
661 TGACACTTGGGCCCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATATG 720

QY 721 TTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCTCTGTGTTGTGA 780
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
721 TTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCTCTGTGTTGTGA 780

QY 781 CTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTCCCAAAGATCTCCCCAT 840
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
781 CTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTCCCAAAGATCTCCCCAT 840

QY 841 TATGTTTGATGTCCTGATTCATGATCCTAGCCACTTCTCAATTATTCTACCAATTAAC 900
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
841 TATGTTTGATGTCCTGATTCATGATCCTAGCCACTTCTCAATTATTCTACCAATTAAC 900

QY 901 CAAGTGGAGCTTCGGGGATAATACTAGCCTGTTTGTTCACCAATCATACTGTGAATCA 960
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
901 CAAGTGGAGCTTCGGGGATAATACTAGCCTGTTTGTTCACCAATCATACTGTGAATCA 960

QY 961 CACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCAGG 1020
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
961 CACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCAGG 1020

QY 1021 ACCTTGTCCGCCACCGCCACCAACCCAGACCTTCAAAAACCCACCCCTTCTTTAGGACC 1080
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1021 ACCTTGTCCGCCACCGCCACCAACCCAGACCTTCAAAAACCCACCCCTTCTTTAGGACC 1080

QY 1081 TGCTGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTAAACAG 1140
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1081 TGCTGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTAAACAG 1140

QY 1141 ATATGGCCACTTTCAGGCCACCATCACAATTTGTAGAGGGAATCTTAGAGGTTAACATCAT 1200
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1141 ATATGGCCACTTTCAGGCCACCATCACAATTTGTAGAGGGAATCTTAGAGGTTAACATCAT 1200

QY 1201 CCAGATGACAGACGTCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTGT 1260
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1201 CCAGATGACAGACGTCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTTGT 1260

QY 1261 CGTGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACCATCATTTCTGACCCACCTG 1320
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1261 CGTGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACCATCATTTCTGACCCACCTG 1320

QY 1321 CGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTGAC 1380
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1321 CGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTGAC 1380

QY 1381 TGTGAGACGAACTTCAATGGGTCTGGAGCTACTGTGTGAACCTCACCTGGGGGATGA 1440
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1381 TGTGAGACGAACTTCAATGGGTCTGGAGCTACTGTGTGAACCTCACCTGGGGGATGA 1440

QY 1441 CACAAGCCTGGCTCTCAGAGCACCCCTGATTTCTGTTCCCTGACAGAGACCCAGCCTCGCC 1500
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1441 CACAAGCCTGGCTCTCAGAGCACCCCTGATTTCTGTTCCCTGACAGAGACCCAGCCTCGCC 1500

QY 1501 TTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACTGT 1560
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1501 TTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACTGT 1560

QY 1561 GATCTCCCTCTTGGTGTACAAAAACACAAAGGAATACACCCCAATAGAAAAATAGTCTGG 1620
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1561 GATCTCCCTCTTGGTGTACAAAAACACAAAGGAATACACCCCAATAGAAAAATAGTCTGG 1620

QY 1621 GAATGTGGTCAGAAAGCAAGGCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTTCTT 1680
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1621 GAATGTGGTCAGAAAGCAAGGCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTTCTT 1680

QY 1681 CCGGGAAACCCAGGAAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAGGAGTTCTTA 1740
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1681 CCGGGAAACCCAGGAAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAGGAGTTCTTA 1740

QY 1741 AATTTCGACCTTGTCTGAAGCTCACCTTTTCAGTGCCCATTTGATGTGAGATGTGCTGGAG 1800
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1741 AATTTCGACCTTGTCTGAAGCTCACCTTTTCAGTGCCCATTTGATGTGAGATGTGCTGGAG 1800

QY 1801 TGGCTATTAAACCTTTTCTCTAAAGATTATTGTTAAATAGATATTGTGTTTGGGGAAG 1860
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1801 TGGCTATTAAACCTTTTCTCTAAAGATTATTGTTAAATAGATATTGTGTTTGGGGAAG 1860

QY 1861 TTGAATTTTATAGGTTAAATGTCAATTTTAGAGATGGGAGAGGATTATATCTGCAGGC 1920
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1861 TTGAATTTTATAGGTTAAATGTCAATTTTAGAGATGGGAGAGGATTATATCTGCAGGC 1920

QY 1921 AGCTTCAGCCATGTTGTGAAACTGTATAAAAGCACTTAGCAAGGCTTCTTTTCATTATT 1980
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1921 AGCTTCAGCCATGTTGTGAAACTGTATAAAAGCACTTAGCAAGGCTTCTTTTCATTATT 1980

QY 1981 TTTATGTTTCACCTTATAAAGTCTTAGGTAACCTAGTAGATAGAAACACACTGTGTCCCGAGA 2040
Db |||||
QY 1981 TTTATGTTTCACCTTATAAAGTCTTAGGTAACCTAGTAGATAGAAACACACTGTGTCCCGAGA 2040
Db |||||
QY 2041 GTAAGGAGAGAAGCTACTATTGATTAGAGCCTAACCCAGGTTAACTGCAAGAAGAGCGG 2100
Db |||||
QY 2041 GTAAGGAGAGAAGCTACTATTGATTAGAGCCTAACCCAGGTTAACTGCAAGAAGAGCGG 2100
Db |||||
QY 2101 GATACCTTCAGCTTTCCATGTAACTGTATGCATAAAGCCAAATGTAGTCCAGTTTCTAAGA 2160
Db |||||
QY 2101 GATACCTTCAGCTTTCCATGTAACTGTATGCATAAAGCCAAATGTAGTCCAGTTTCTAAGA 2160
Db |||||
QY 2161 TCATGTTCCAAAGCTAACTGAATCCCACCTTCAATACACACTCATGAACCTCCTGATGGAACA 2220
Db |||||
QY 2161 TCATGTTCCAAAGCTAACTGAATCCCACCTTCAATACACACTCATGAACCTCCTGATGGAACA 2220
Db |||||
QY 2221 ATAAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGCTAGACTCAGAAAAAATACTAC 2280
Db |||||
QY 2221 ATAAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGCTAGACTCAGAAAAAATACTAC 2280
Db |||||
QY 2281 TCTCATAAATGGTGGGAGTATTTGGTGTGACAACCTACTTTGCTGGCTGAGTGAAGGAA 2340
Db |||||
QY 2281 TCTCATAAATGGTGGGAGTATTTGGTGTGACAACCTACTTTGCTGGCTGAGTGAAGGAA 2340
Db |||||
QY 2341 TGATATTCATATATTCAATTTATCCATGGACATTTAGTTAGTGTCTTTTATATACCAGGC 2400
Db |||||
QY 2341 TGATATTCATATATTCAATTTATCCATGGACATTTAGTTAGTGTCTTTTATATACCAGGC 2400
Db |||||
QY 2401 ATGATGCTGAGTGACACTCTGTGTATATTTCCAAATTTTGTATAGTCGCTGCACATAT 2460
Db |||||
QY 2401 ATGATGCTGAGTGACACTCTGTGTATATTTCCAAATTTTGTATAGTCGCTGCACATAT 2460
Db |||||
QY 2461 TTGAAATCAAAATATTAAAGACTTTCCAAAAATTTGGTCCTCGTGTTCATGGCAACTTG 2520
Db |||||
QY 2461 TTGAAATCAAAATATTAAAGACTTTCCAAAAATTTGGTCCTCGTGTTCATGGCAACTTG 2520
Db |||||
QY 2521 ATCAGTAAGGATTTCCCTCTGTTTGGAACTAAACCAATTTACTATATGTTAGACAAGAC 2580
Db |||||
QY 2521 ATCAGTAAGGATTTCCCTCTGTTTGGAACTAAACCAATTTACTATATGTTAGACAAGAC 2580
Db |||||
QY 2581 ATTTTCTCTCTGAAAGAAAAAATGAGGGAAGAGACAAAAA 2640
Db |||||
QY 2581 ATTTTCTCTGAAAGAAAAAATGAGGGAAGAGACAAAAA 2640
Db |||||
QY 2641 AAAAAA 2661
Db |||||
QY 2641 AAAAAA 2661
Db |||||

RESULT 2
ID AAT69328 standard; cDNA; 2669 BP.
XX AAT69328;
AC AAT69328;
XX
DT 26-FEB-1998 (first entry)
XX
DE Murine metastatic nucleic acid sequence.
XX
KW Mouse; murine; tumour; cancer; metastatic sequence; detection; diagnosis;
KW treatment; metastasis; hyperplasia; dysplasia; hypertrophy; screening;
KW ss.
XX
OS Mus musculus.
XX
PN WO9718454-A2.
XX
PD 22-MAY-1997.
XX
PF 15-NOV-1996; 96WO-US018567.
XX
PR 16-NOV-1995; 95US-0006838P.
PR 30-JAN-1996; 96US-00594031.

XX
PA
XX (THOM/) THOMPSON T.
PI Thompson T;
XX
DR WPI; 1997-289397/26.
XX
PT Identifying tumour metastatic sequences - by introducing transfected
PT cells into host mammal and analysing primary and metastatic sequences by
PT differential display PCR.
XX
PS Disclosure; Fig 12CS; 102pp; English.
XX
CC Mouse Urogenital Sinus (UGS) tissue was isolated from 17 day old mouse
CC embryos. The UGS cells were infected with retroviruses, cultured and
CC implanted under the renal capsule of mice. Reconstitutions were harvested
CC 5 weeks later, when they showed signs of distress from the tumour burden.
CC Metastatised tumours were isolated from a site outside the renal capsule.
CC RNA was isolated from primary tumours and metastases, reverse transcribed
CC and subjected to differential display PCR. The sequences were analysed to
CC obtain metastatic sequences, e.g. the present sequence. The method can be
CC used to detect, diagnose and treat disorders related to metastasis, or
CC treat malignant or non-malignant disorders, e.g. hyperplasia, dysplasia
CC and hypertrophy. The metastatic sequence can be used to screen a
CC biological sample for metastasis, and it or its expression product may
CC also be used to treat a metastatic disorder
XX
SQ Sequence 2669 BP; 752 A; 595 C; 597 G; 725 T; 0 U; 0 Other;
Query Match 95.7%; Score 2547.8; DB 2; Length 2669;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;
QY 28 AACCTTGGTGCCCTGCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTCCTGGGAT 87
Db |||||
QY 60 ACCTTGAGTGCCCTGCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTCCTGGGAT 119
Db |||||
QY 88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCGGCCCAACGAATTCATGATGTCTGG 147
Db |||||
QY 120 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCGGCCCAACGAATTCATGATGTCTGG 179
Db |||||
QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGCTTCTG 207
Db |||||
QY 180 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGCTTCTG 239
Db |||||
QY 208 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 267
Db |||||
QY 240 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 299
Db |||||
QY 268 AAAACTCCTGGAAGGGAGGCCGTGTGCGGTGAACCTGATATTCCTTAGATGCCAAAGGAAGATG 327
Db |||||
QY 300 AAAACTCCTGGAAGGGAGGCCGTGTGCGGTGAACCTGATATTCCTTAGATGCCAAAGGAAGATG 359
Db |||||
QY 328 TGGGCTCAATATAAATTTGCGGTGAACCTGATATTCCTTAGATGCCAAAGGAAGATG 387
Db |||||
QY 360 TGGGCTCAATATAAATTTGCGGTGAACCTGATATTCCTTAGATGCCAAAGGAAGATG 419
Db |||||
QY 388 CCAATGGCAACATAGTCTATGAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 447
Db |||||
QY 420 CCAATGGCAACATAGTCTATGAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 479
Db |||||
QY 448 CATATGTTTACAACTGGACAGATGGTGCAGAGGACAGTGACGGGAAAAATGGCACC GCC 507
Db |||||
QY 480 CATATGTTTACAACTGGACAGATGGTGCAGAGGACAGTGACGGGAAAAATGGCACC GCC 539
Db |||||
QY 508 AAAGCCATCAACGCTCTTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAA 567
Db |||||
QY 540 AAAGCCATCAACGCTCTTCCCTGATGGGAAACCTTTTCTCACCACCCCGGATGGAGAA 599
Db |||||
QY 568 GATGGAATTTCACTACGCTCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT 627
Db |||||
QY 600 GATGGAATTTCACTACGCTCTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT 659
Db |||||

XX Mouse; murine; tumour; cancer; metastatic sequence; detection; diagnosis;
KW treatment; metastasis; hyperplasia; dysplasia; hypertrophy; screening;
KW ss.
XX Mus musculus.
XX PN WO9718454-A2.
XX PD 22-MAY-1997.
XX PF 15-NOV-1996; 96WO-US018567.
XX PR 16-NOV-1995; 95US-0006838P.
XX PR 30-JAN-1996; 96US-00594031.
XX (THOM/) THOMPSON T.
XX PA Thompson T;
XX PI
XX DR WPI; 1997-289397/26.
XX PT Identifying tumour metastatic sequences - by introducing transfected
PT cells into host mammal and analysing primary and metastatic sequences by
PT differential display PCR.
XX PS Disclosure; Fig 12BT; 102pp; English.
XX CC Mouse Urogenital Sinus (UGS) tissue was isolated from 17 day old mouse
CC embryos. The UGS cells were infected with retroviruses, cultured and
CC implanted under the renal capsule of mice. Reconstitutions were harvested
CC 5 weeks later, when they showed signs of distress from the tumour burden.
CC Metastasised tumours were isolated from a site outside the renal capsule.
CC RNA was isolated from primary tumours and metastases, reverse transcribed
CC and subjected to differential display PCR. The sequences were analysed to
CC obtain metastatic sequences, e.g. the present sequence. The method can be
CC used to detect, diagnose and treat disorders related to metastasis, or
CC treat malignant or non-malignant disorders, e.g. hyperplasia, dysplasia
CC and hypertrophy. The metastatic sequence can be used to screen a
CC biological sample for metastasis, and it or its expression product may
CC also be used to treat a metastatic disorder
xx
SQ Sequence 2669 BP; 752 A; 595 C; 597 G; 725 T; 0 U; 0 Other;

Query Match 95.7%; Score 2547.8; DB 2; Length 2669;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;
QY 28 AACCTTGGTGCCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCTGGGAT 87
Db 60 ACCTTGAGTGCCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCTGGGAT 119
QY 88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTTCATGATGTCTGG 147
Db 120 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTTCATGATGTCTGG 179
QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTAAATGGCTGGTCTTCTG 207
Db 180 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTAAATGGCTGGTCTTCTG 239
QY 208 ATGAAAATGACTGGAATGAAAACTCTACCCAGTGTGAAGCGGGAGACATGAGGTGGA 267
Db 240 ATGAAAATGACTGGAATGAAAACTCTACCCAGTGTGAAGCGGGAGACATGAGGTGGA 299
QY 268 AAAACTCCTGGAGGGAGGCCGTGTGCAGGCGGTCTGACCCAGTCACTACCCAGCCCTCG 327
Db 300 AAAACTCCTGGAGGGAGGCCGTGTGCAGGCGGTCTGACCCAGTCACTACCCAGCCCTCG 359
QY 328 TGGGCTCAATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAAGGAAGT 387
Db 360 TGGGCTCAATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAAGGAAGT 419
QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 447

Db 420 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 479
QY 448 CATATGTTTACAACTGGACAGCATGGTTCAGAGACAGTACGGGAAAAATGGACCCGGCC 507
Db 480 CATATGTTTACAACTGGACAGCATGGTTCAGAGACAGTACGGGAAAAATGGACCCGGCC 539
QY 508 AAAGCCATCATAAACGTCTTCCCTGATGGGAAACCTTTTCTCCACACCCCGGATGGAGAA 567
Db 540 AAAGCCATCATAAACGTCTTCCCTGATGGGAAACCTTTTCTCCACACCCCGGATGGAGAA 599
QY 568 GATGGAATTTTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGAT 627
Db 600 GATGGAATTTTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGACGAT 659
QY 628 GTTCAGTGAGAGTTTCTGTGAACACAGACCAATGTGACACTTTGGGCTCAACTCATGGAAG 687
Db 660 GTTCAGTGAGAGTTTCTGTGAACACAGACCAATGTGACACTTTGGGCTCAACTCATGGAAG 719
QY 688 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 747
Db 720 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 779
QY 748 ACGTGGTAAACAGATCAGATTCTGTGTTTGTGACTATGTTCCAGAAAGAACGATCGAAATT 807
Db 780 ACGTGGTAAACAGATCAGATTCTGTGTTTGTGACTATGTTCCAGAAAGAACGATCGAAATT 839
QY 808 CATCCGACGAAACCTTCC-CAAAAGATCTCCCCATTTATGTTGATGTCTGATTCATGATC 866
Db 840 CATCCGACGAAACCTTCCCTCAAAGATCTCCCCATTTATGTTGATGTCTGATTCATGATC 899
QY 867 CTAGCCACTTCTCAATTTTCTACCAATTAACACTACAGTGGAGCTTCGGGGATAAATCTG 926
Db 900 CTAGCCACTTCTCAATTTTCTACCAATTAACACTACAGTGGAGCTTCGGGGATAAATCTG 959
QY 927 GCCTGTTTGTTCACCAATCATACTGTGAATCACACGATGTCTCAATGGAACCTTCA 986
Db 960 GCCTGTTTGTTCACCAATCATACTGTGAATCACACGATGTCTCAATGGAACCTTCA 1019
QY 987 GCCTTAACCTCACTGTGAAAGCTGCAGCACCCAGGACCTTGTCCGCCACCCGCCACCCAC 1046
Db 1020 GCCTTAACCTCACTGTGAAAGCTGCAGCACCCAGGACCTTGTCCGCCACCCGCCACCCAC 1079
QY 1047 CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGTGTTGACAAACCCCTGGAGCTGA 1106
Db 1080 CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGTGTTGACAAACCCCTGGAGCTGA 1139
QY 1107 GTAGGATTCCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCA 1166
Db 1140 GTAGGATTCCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCA 1199
QY 1167 CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCTGATGCCGG 1226
Db 1200 CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCTGATGCCGG 1259
QY 1227 TGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGTCGACCTGCCAAGGGAGCATTCCTCA 1286
Db 1260 TGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGTCGACCTGCCAAGGGAGCATTCCTCA 1319
QY 1287 CGGAGGTCTGTAACCATCATTTCTGACCCCACTCGAGATCACCCAGAACACAGTCTGCA 1346
Db 1320 CGGAGGTCTGTACCATCATTTCTGACCCCACTCGAGATCACCCAGAACACAGTCTGCA 1379
QY 1347 GCCCTGTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGAACTTCAATGGGTCTG 1406
Db 1380 GCCCTGTGGATGTGGATGAGATGTGTCTGCTGACTGTGAGACGAACTTCAATGGGTCTG 1439
QY 1407 GGACGTAATGTGTGAACCTCACCCCTGGGGATGACACAAGCCTGGCTCTCACGAGCACCC 1466
Db 1440 GGACGTAATGTGTGAACCTCACCCCTGGGGATGACACAAGCCTGGCTCTCACGAGCACCC 1499
QY 1467 TGATTTCTGTTCTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCTGA 1526

Db 1500 TGATTTCTGTTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCCTGA 1559

QY 1527 TCTCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGACAAAAAC 1586

Db 1560 TCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGACAAAAAC 1619

QY 1587 ACAAGGAATACAAACCCCAATAGAAAATAGTCCCTGGGAATGTGGTCAGAAAGCAAGGCCTGA 1646

Db 1620 ACAAGGAATACAAACCCCAATAGAAAATAGTCCCTGGGAATGTGGTCAGAAAGCAAGGCCTGA 1679

QY 1647 GTGTCTTTCTCAACCGTGCAAAAGCCGTTCTTCCCGGAAACCCAGGAAAGGATCCGC 1706

Db 1680 GTGTCTTTCTCAACCGTGCAAAAGCCGTTCTTCCCGGAAACCCAGGAAAGGATCCGC 1739

QY 1707 TACTCAAAAACCAAGAAATTTAAAGGAGTTTCTTAAATTTTCGACCTTGTCTGAAGCTCA 1766

Db 1740 TACTCAAAAACCAAGAAATTTAAAGGAGTTTCTTAAATTTTCGACCTTGTCTGAAGCTCA 1799

QY 1767 CTTTTCAGTGCCATTGATGTGAGATGTCTGAGTGGCTATTAAACCTTTTTTTCCTAAAG 1826

Db 1800 CTTTTCAGTGCCATTGATGTGAGATGTCTGAGTGGCTATTAAACCTTTTTTTCCTAAAG 1859

QY 1827 ATTATTGTTAAATAGATATTGTGTTGGGGAAGTTGAATTTTTTATAGTTAAATGTCA 1886

Db 1860 ATTATTGTTAAATAGATATTGTGTTGGGGAAGTTGAATTTTTTATAGTTAAATGTCA 1919

QY 1887 TTTTAGAGATGGGAGAGGGATTATATCTGCAGGAGCTTCAGCCATGTTGTGAAACTGAT 1946

Db 1920 TTTTAGAGATGGGAGAGGGATTATATCTGCAGGAGCTTCAGCCATGTTGTGAAACTGAT 1979

QY 1947 AAAAGCAACTTAGCAAGGCTTCTTTTTCATTATTTTATGTTTCACTTATAAAGCTTAG 2006

Db 1980 AAAAGCAACTTAGCAAGGCTTCTTTTTCATTATTTTATGTTTCACTTATAAAGCTTAG 2039

QY 2007 GTAACCTAGTAGAATAACACTGTGTCCCGAGAGTAAGGAGAGAGCTACTATTGATTA 2066

Db 2040 GTAACCTAGTAGAATAACACTGTGTCCCGAGAGTAAGGAGAGAGCTACTATTGATTA 2099

QY 2067 GAGCCTAACCCAGGTTAACTGCAAGAAGAGCGGGATACTTTTCAGCTTTCCATGTAACGTG 2126

Db 2100 GAGCCTAACCCAGGTTAACTGCAAGAAGAGCGGGATACTTTTCAGCTTTCCATGTAACGTG 2159

QY 2127 TATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTGTCCCAAGCTAACTGAATCCCA 2186

Db 2160 TATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTGTCCCAAGCTAACTGAATCCCA 2219

QY 2187 CTTCAATACACACTCATGAACCTCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGA 2246

Db 2220 CTTCAATACACACTCATGAACCTCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGA 2279

QY 2247 TGTGCACACTTGTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTATTCCCA 2306

Db 2280 TGTGCACACTTGTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTATTCCCA 2339

QY 2307 GTGACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTATTCCCA 2366

Db 2340 GTGACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTATTCCCA 2399

QY 2367 TGGACATTTAGTTAGTGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTTGTGTA 2426

Db 2400 TGGACATTTAGTTAGTGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTTGTGTA 2459

QY 2427 TATTTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAAATATTAAGACTTTCC 2486

Db 2460 TATTTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATC-ATATATTAAGACTTTCC 2518

QY 2487 AAAAATTTGGTCCCTGGTTTTCATGGCACTTGATCAGTAAGGATTTCCCTCTGTTTG 2546

Db 2519 AAAGATGAGTCCCTGGTTTTCATGGCACTTGATCAGTAAGGATTTTCACCTCTGTTTG 2578

QY 2547 GAACATAAACCATTTTACTATATGTTAGACAAGACATTTTTCCTTCTCTGAAAAA 2606

Db 2579 TAACTAAACCATCTACTATATGTTAGACATGACATTTCTTTCTCTCTCTCTGAAAAA 2638

QY 2607 -AAAAATGAGGGAAGAGACAAAAAAGAAAAA 2636

Db 2639 TAAAGTGTGGAGAGACAAAAAAGAAAAA 2669

RESULT 4
AAC55715
ID AAC55715 standard; cDNA; 2669 BP.
XX
AC AAC55715;
XX 17-JAN-2001 (first entry)
XX Human NMB cDNA sequence from Genbank X76534.
DE Human; differentially regulated gene; macrophage development; diagnosis;
XX matrix metalloproteinase 19; MMP19; antiarthritic; antiinflammatory;
KW destructive macrophage development inhibitor; arthritis;
KW colorectal cancer; immune response; ss.
XX Homo sapiens.
XX WO200055373-A2.
XX 21-SEP-2000.
XX 15-MAR-2000; 2000WO-US006883.
XX 15-MAR-1999; 99US-0124530P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX Murray R;
XX WPI; 2000-628200/60.
XX Screening drug candidates comprises adding a drug to a cell expressing an
expression profile gene and determining the effect of the drug on the
expression of the expression profile gene.
XX Claim 1; Page; 99pp; English.
XX The present invention describes a method for screening drug candidates.
CC The method comprises adding a drug to a cell that expresses an expression
CC profile gene encoding a protein encoded by 5 sequences of defined base
CC pairs as given in C55638, C55642, C55643, C55644 and C55653 or a sequence
CC represented by Genbank accession number X92521, X62466, J04130, X62087
CC and X76534 (or a fragment) and determining the effect of the drug on the
CC expression of the expression profile gene. An inhibitor of matrix
CC metalloproteinase 19 (MMP-19), preferably an antibody, is useful for
CC treating destructive macrophage disorders (DMD) by inhibiting DM
CC development in a cell of an individual having arthritis. Antibodies to
CC MMP-19 are useful for localising a therapeutic moiety preferably
CC cytotoxic agent or a radioisotope to colorectal cancer tissue. A
CC composition comprising MMP-19 is useful for eliciting an immune response
CC in an individual. C55635 to C55710 represent human differentially
CC regulated genes of the invention. The present sequence represents the
CC human NMB cDNA sequence according to the Genbank accession number X76534.
CC N.B. The present sequence is not given in the present specification, but
CC it is specifically claimed by its Genbank accession number
XX
SQ Sequence 2669 BP; 752 A; 595 C; 597 G; 725 T; 0 U; 0 Other;
Query Match 95.7%; Score 2547.8; DB 3; Length 2669;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 28 AACCTTGGTGGCTCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCTCTGGGAT 87

Db 60 ACCTTGAGTGGCTCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCTCTGGGAT 119

QY 88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCCAACGATTCATGATGTGCTGG 147

Db 120 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCGGCCAAACGATTTTCATGATGCTGG 179
QY 148 GCAATGAAAGACTTCTGCTTACATGAGGAGCACAATCAATTAATGGCTGGTCTTCTG 207
Db 180 GCAATGAAAGACTTCTGCTTACATGAGGAGCACAATCAATTAATGGCTGGTCTTCTG 239
QY 208 ATGAAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGA 267
Db 240 ATGAAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGA 299
QY 268 AAAACTCCTGGAAGGAGGCCGCTGTCAGGCGGTCTGACCGAGTGACTCACAGCCCTCG 327
Db 300 AAAACTCCTGGAAGGAGGCCGCTGTCAGGCGGTCTGACCGAGTGACTCACAGCCCTCG 359
QY 328 TGGGCTCAAATATAACATTTGCGGTGAACCTGATATTCCTTAGATGCCAAAAGGAAGTG 387
Db 360 TGGGCTCAAATATAACATTTGCGGTGAACCTGATATTCCTTAGATGCCAAAAGGAAGTG 419
QY 388 CCAATGSCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 447
Db 420 CCAATGSCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 479
QY 448 CATATGTTTACAACCTGGACAGCATGGTCAGAGGACAGTGACGGGAAAAATGGCACCCGCC 507
Db 480 CATATGTTTACAACCTGGACAGCATGGTCAGAGGACAGTGACGGGAAAAATGGCACCCGCC 539
QY 508 AAAGCCATCATAAACGTCTTCCCTGATGGGAAACCTTTTCCACACACCCCGGATGGAGAA 567
Db 540 AAAGCCATCATAAACGTCTTCCCTGATGGGAAACCTTTTCCACACACCCCGGATGGAGAA 599
QY 568 GATGGAATTTTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT 627
Db 600 GATGGAATTTTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT 659
QY 628 GTTCAGTGAGAGTTTCTGTGAACACAGCCAAATGTGACACTTGGGCCTCAACTCATGGAAG 687
Db 660 GTTCAGTGAGAGTTTCTGTGAACACAGCCAAATGTGACACTTGGGCCTCAACTCATGGAAG 719
QY 688 TGA CTGTCTACAGAAAGACATGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 747
Db 720 TGA CTGTCTACAGAAAGACATGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 779
QY 748 ACGTGGTAAACAGATCAGATTCCTGTGTTGTGACTATCTCCCATGATGTTTATGATTCATGATC 807
Db 780 ACGTGGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAACGATCGAAAT 839
QY 808 CATCCGACGAAACCTTCC - CAAAGATCTCCCATTAATGTTTCCAGAAACGATCGAAAT 866
Db 840 CATCCGACGAAACCTTCCCATTAATGTTTCCCATTAATGTTTCCAGAAACGATCGAAAT 899
QY 867 CTAGCCACTTCCATTAATTTCTACCATTAACCTAAGTGGAGCTTCGGGGATAACTG 926
Db 900 CTAGCCACTTCCATTAATTTCTACCATTAACCTAAGTGGAGCTTCGGGGATAACTG 959
QY 927 GCCTGTTTGTTCACCAATCATACTGTGAATCACACGTATGTGCTCAATGGAACCTTCA 986
Db 960 GCCTGTTTGTTCACCAATCATACTGTGAATCACACGTATGTGCTCAATGGAACCTTCA 1019
QY 987 GCCTTAACCTCACTGTGAAAGCTGCAGCACGAGACCTTGTCCGCCACCCGACCCAC 1046
Db 1020 GCCTTAACCTCACTGTGAAAGCTGCAGCACGAGACCTTGTCCGCCACCCGACCCAC 1079
QY 1047 CCAGACCTTCAAAACCCACCTTCTTTTAGGACCTGCTGGTGACAAACCCCTGGAGCTGA 1106
Db 1080 CCAGACCTTCAAAACCCACCTTCTTTTAGGACCTGCTGGTGACAAACCCCTGGAGCTGA 1139
QY 1107 GTAGGATTCCTGATGAAAACTGCCAGATTAACAGATATGGCCACTTTCAAGCCACCATCA 1166
Db 1140 GTAGGATTCCTGATGAAAACTGCCAGATTAACAGATATGGCCACTTTCAAGCCACCATCA 1199
QY 1167 CAATTGTAGGGGAATCTTAGAGGTTAACATCATCCAGATGACAGACGCTCCTGATGCCGG 1226

Db 1200 CAATTGTAGAGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATGCCGG 1259
QY 1227 TGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTCTGTGACCTGCCAAGGAGCATTTCCCA 1286
Db 1260 TGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTCTGTGACCTGCCAAGGAGCATTTCCCA 1319
QY 1287 CGGAGGTCTGTACCATCATTTCTGACCCACCTGCGAGATCACCCAGAAACACAGTCTGCA 1346
Db 1320 CGGAGGTCTGTACCATCATTTCTGACCCACCTGCGAGATCACCCAGAAACACAGTCTGCA 1379
QY 1347 GCCCTGTGGATGTGGATGAGATGTCTGTCTGACTGTGAGACGAACTTCAATGGGTCTG 1406
Db 1380 GCCCTGTGGATGTGGATGAGATGTCTGTCTGACTGTGAGACGAACTTCAATGGGTCTG 1439
QY 1407 GGACGTACTGTGGAACCTCACCCCTGGGGGATGACACAAGCCTGGCTCACGAGCACCC 1466
Db 1440 GGACGTACTGTGGAACCTCACCCCTGGGGGATGACACAAGCCTGGCTCACGAGCACCC 1499
QY 1467 TGATTTCTGTTCTTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTCCCTGA 1526
Db 1500 TGATTTCTGTTCTTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTCCCTGA 1559
QY 1527 TCTCCGTTGGCTGCTTGCCCAATATTGTCACTGTGATCTCCCTCTTGGTGACAAAAAC 1586
Db 1560 TCTCCGTTGGCTGCTTGCCCAATATTGTCACTGTGATCTCCCTCTTGGTGACAAAAAC 1619
QY 1587 ACAAGGAATACAAACCCCAATAGAAAAATAGTCTCTGGGAATGTGGTCAGAAAGCAAGCCTGA 1646
Db 1620 ACAAGGAATACAAACCCCAATAGAAAAATAGTCTCTGGGAATGTGGTCAGAAAGCAAGCCTGA 1679
QY 1647 GTGTCTTTCTCAACCCGTGCAAAAGCCGTGTTCTTCCGGGAAACAGGAAAGGATCCGC 1706
Db 1680 GTGTCTTTCTCAACCCGTGCAAAAGCCGTGTTCTTCCGGGAAACAGGAAAGGATCCGC 1739
QY 1707 TACTCAAAAACCAAGATTTAAAGGAGTTTCTTAAATTCGACCTTGTCTGAGCTCA 1766
Db 1740 TACTCAAAAACCAAGATTTAAAGGAGTTTCTTAAATTCGACCTTGTCTGAGCTCA 1799
QY 1767 CTTTTCAGTGCCATTGATGTGAGATGTGCTGGAGTGCCTATTAACCTTTTCTCTAAAG 1826
Db 1800 CTTTTCAGTGCCATTGATGTGAGATGTGCTGGAGTGCCTATTAACCTTTTCTCTAAAG 1859
QY 1827 ATTATTGTTAAATAGATATTGTGTTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA 1886
Db 1860 ATTATTGTTAAATAGATATTGTGTTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA 1919
QY 1887 TTTTAGAGATGGGAGAGGGATTATACTGCAGGAGCTTCAGCCATGTTGTGAAACTGAT 1946
Db 1920 TTTTAGAGATGGGAGAGGGATTATACTGCAGGAGCTTCAGCCATGTTGTGAAACTGAT 1979
QY 1947 AAAAGCAACTTAGCAAGGCTTCTTTTCATTAATTTTATGTTTCACTTATAAAGTCTTAG 2006
Db 1980 AAAAGCAACTTAGCAAGGCTTCTTTTCATTAATTTTATGTTTCACTTATAAAGTCTTAG 2039
QY 2007 GTAACTAGTAGGATAGAAAACACTGTGTCCGAGAGTAAGGAGAGAGCTACTATTGATTA 2066
Db 2040 GTAACTAGTAGGATAGAAAACACTGTGTCCGAGAGTAAGGAGAGAGCTACTATTGATTA 2099
QY 2067 GAGCCTAACCCAGGTTAACTGCAAGAAAGAGCGGGATACTTTTCAGCTTTCCATGTAAGT 2126
Db 2100 GAGCCTAACCCAGGTTAACTGCAAGAAAGAGCGGGATACTTTTCAGCTTTCCATGTAAGT 2159
QY 2127 TATGCATAAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAGCTAACTGAATCCCA 2186
Db 2160 TATGCATAAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAGCTAACTGAATCCCA 2219
QY 2187 CTTCAATACACACTCATGAACCTCCTGATGGAAACAATAACAGGCCCAAGCCTGTGGTATGA 2246
Db 2220 CTTCAATACACACTCATGAACCTCCTGATGGAAACAATAACAGGCCCAAGCCTGTGGTATGA 2279
QY 2247 TGTGCACACTTGCTAGACTCAGAAAAATACTACTCTCATAAATGGGTGGGAGTATTTTG 2306
Db 2280 TGTGCACACTTGCTAGACTCAGAAAAATACTACTCTCATAAATGGGTGGGAGTATTTTG 2339

CC induced abnormalities in bone formation or bone loss, conditions that
CC involve altered bone metabolism (e.g. idiopathic juvenile osteoporosis),
CC skeletal disease linked to breast cancer, mastocytosis, Fanconi syndrome
CC or fibrous dysplasia. The present sequence is that of an osteoblast
CC differentiation associated cDNA marker of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2669 BP; 752 A; 595 C; 597 G; 725 T; 0 U; 0 Other;

Query Match 95.7%; Score 2547.8; DB 6; Length 2669;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 28 AACCTTGGTGGCTGCGTCCGTHGAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGGAT 87
Db |||||
QY 60 ACCTTGAGTGGCTGCGTCCGTHGAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGGAT 119
Db |||||
QY 88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTTCATGATGTCTGG 147
Db |||||
QY 120 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTTCATGATGTCTGG 179
Db |||||
QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGAGCACAATCAATTAATGGCTGGTCTTCTG 207
Db |||||
QY 180 GCAATGAAAGACCTTCTGCTTACATGAGGAGCACAATCAATTAATGGCTGGTCTTCTG 239
Db |||||
QY 208 ATGAAATGACTGGAAATGAAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 267
Db |||||
QY 240 ATGAAATGACTGGAAATGAAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 299
Db |||||
QY 268 AAAACTCCTGGAGAGGAGGCCGTGTGCAGGGCTCTGACCACTGACTCACCAGCCCTCG 327
Db |||||
QY 300 AAAACTCCTGGAGAGGAGGCCGTGTGCAGGGCTCTGACCACTGACTCACCAGCCCTCG 359
Db |||||
QY 328 TGGGCTCAAAATATAACATTTGCGGTGAACCTGTATATTCCTAGATGCCAAAAGAGATG 387
Db |||||
QY 360 TGGGCTCAAAATATAACATTTGCGGTGAACCTGTATATTCCTAGATGCCAAAAGAGATG 419
Db |||||
QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 447
Db |||||
QY 420 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 479
Db |||||
QY 448 CATATGTTTACAACCTGGACAGCATGGTTCAGAGGACACTGACGGGAAAATGGCACCGGCC 507
Db |||||
QY 480 CATATGTTTACAACCTGGACAGCATGGTTCAGAGGACACTGACGGGAAAATGGCACCGGCC 539
Db |||||
QY 508 AAAGCCATCATAACGCTTCCCTGATGGGAAACCTTTCTCACCACCCCGGATGGAGAA 567
Db |||||
QY 540 AAAGCCATCATAACGCTTCCCTGATGGGAAACCTTTCTCACCACCCCGGATGGAGAA 599
Db |||||
QY 568 GATGGAATTTTCATCTACGCTTCCACACACTTGGTCAAGTATTTCCAGAAATTTGGGACGAT 627
Db |||||
QY 600 GATGGAATTTTCATCTACGCTTCCACACACTTGGTCAAGTATTTCCAGAAATTTGGGACGAT 659
Db |||||
QY 628 GTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG 687
Db |||||
QY 660 GTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG 719
Db |||||
QY 688 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAGTGAAGATGTGT 747
Db |||||
QY 720 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAGTGAAGATGTGT 779
Db |||||
QY 748 ACGTGGTAAACAGATCAGATTCTCTGTGTTTGTGACTATGTTCCAGAAAGAACGATCGAAAT 807
Db |||||
QY 780 ACGTGGTAAACAGATCAGATTCTCTGTGTTTGTGACTATGTTCCAGAAAGAACGATCGAAAT 839
Db |||||
QY 808 CATCCGACGAAACCTTCC-CAAGATCTCCCCATTATGTTTGTATGTTCTGATTCATGATC 866
Db |||||
QY 840 CATCCGACGAAACCTTCTCTCAAGATCTCCCCATTATGTTTGTATGTTCTGATTCATGATC 899
Db |||||
QY 867 CTAGCCACTTCTCAATATTCTACATTAACATAAGTGGAGCTTCGGGGATAATCTG 926
Db |||||

QY 2307 GTGACAACTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATTCCA 2366
Db |||||
QY 2340 GTGACAACTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATTCCA 2399
Db |||||
QY 2367 TGGACATTTAGTGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTGTGTA 2426
Db |||||
QY 2400 TGGACATTTAGTGTCTTTTATATACAGGCATGATGCTGAGTGACACTCTTGTGTA 2459
Db |||||
QY 2427 TATTTCCAAATTTTGTATAGTGGCTGCACATATTTGAAATCAAAATATTAAAGACTTTCC 2486
Db |||||
QY 2460 TATTTCCAAATTTTGTATAGTGGCTGCACATATTTGAAATC-ATATATTAAAGACTTTCC 2518
Db |||||
QY 2487 AAAAATTTGGTCCCTGGTTTTTTCATGGAACCTTGATCAGTAAGGATTTCCCTCTGTTTG 2546
Db |||||
QY 2519 AAAGATGAGTCCCTGGTTTTTTCATGGAACCTTGATCAGTAAGGATTTCACTCTGTTTG 2578
Db |||||
QY 2547 GAACTAAACCAATTTACTATATGTTTAGACAAGACATTTTTTTTTTCTCTCTGAAAAA 2606
Db |||||
QY 2579 TAACTAAACCAATCTACTATATGTTTAGACAAGACATTTTTTTTCTCTCTGAAAAA 2638
Db |||||
QY 2607 -AAAATGAGGGAAGACAAAAA 2636
Db |||||
QY 2639 TAAAGTGTGGGAAGACAAAAA 2669
Db |||||

RESULT 5
ABQ88185
ID ABQ88185 standard; cDNA; 2669 BP.
XX
AC ABQ88185;
XX
DT 18-SEP-2002 (first entry)
XX
DE Human osteoblast differentiation related cDNA SEQ ID NO 92.
XX
KW Human; osteoblast; stem cell differentiation; bone tissue deposition;
KW osteoporosis; osteopathic; ss.
XX
OS Homo sapiens.
XX
PN WO200250301-A2.
XX
PD 27-JUN-2002.
XX
PF 18-DEC-2001; 2001WO-US048276.
XX
PR 18-DEC-2000; 2000US-0255882P.
PR 24-APR-2001; 2001US-0285691P.
XX
PA (GENE-) GENE LOGIC INC.
PA (PROC) PROCTER & GAMBLE CO.
XX
PI Ji D, Axelrod DW, Cook JS, Jaiswal N, Einstein R, Houghton A;
PI Mertz L;
XX
DR WPI; 2002-557663/59.
XX
PT Use of genes and their expression profiles associated with osteoblast
PT differentiation for screening modulators bone formation, for diagnosing
PT or treating e.g. osteoporosis, or as markers for the differentiation
PT process.
XX
PS Claim 1; SEQ ID NO 92; 78pp + Sequence Listing; English.
XX
CC The invention relates to genes and their expression profiles are used
CC for: (a) screening modulators of precursor stem cell differentiation into
CC osteoblasts, or bone tissue deposition; (b) diagnosing abnormal
CC deposition of bone tissue, abnormal rate of osteoblast formation or
CC osteoporosis; or (c) treating or monitoring treatment of the conditions
CC cited in (b), or monitoring the progression of bone tissue deposition.
CC Specific conditions include postmenopausal osteoporosis, glucocorticoid
CC osteoporosis or male osteoporosis, osteopenia, osteodystrophy, drug-

Db 900 CTAGCCACTTCCTCAATTATTCTACCAATTAACATAAGTGGAGCTTCGGGGATAATACTG 959

QY 927 GCCTGTTGTTTCCACCAATCAVACTGTGAATCACACGTATGTGCTCAATGGAACCTTCA 986

Db 960 GCCTGTTGTTTCCACCAATCATACTGTGAATCACACGTATGTGCTCAATGGAACCTTCA 1019

QY 987 GCCTTAACCTCACTGTGAAAGCTGCAGCACCAAGACCTTGTCCGCCACCGCCACCACCAC 1046

Db 1020 GCCTTAACCTCACTGTGAAAGCTGCAGCACCAAGACCTTGTCCGCCACCGCCACCACCAC 1079

QY 1047 CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGCTGGTGACAAACCCCTCGGAGCTGA 1106

Db 1080 CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGCTGGTGACAAACCCCTCGGAGCTGA 1139

QY 1107 GTAGGATTCCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCA 1166

Db 1140 GTAGGATTCCTGATGAAAACTGCCAGATTAAACAGATAAGGCCACTTTCAGGCCACCATCA 1199

QY 1167 CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCTCTGATGCCGG 1226

Db 1200 CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCTCTGATGCCGG 1259

QY 1227 TGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGCTGACTGCCAAGGGAGCATTCCTCA 1286

Db 1260 TGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGCTGACTGCCAAGGGAGCATTCCTCA 1319

QY 1287 CGGAGTCTGTACCATCATTTCTGACCCCACTCGGAGATCACCCAGAAACACAGTCTGCA 1346

Db 1320 CGGAGTCTGTACCATCATTTCTGACCCCACTCGGAGATCACCCAGAAACACAGTCTGCA 1379

QY 1347 GCCCTGTGGATGGATGAGATGTGTCTGCTGACTGTGAGACGAAACCTTCAATGGGCTG 1406

Db 1380 GCCCTGTGGATGGATGAGATGTGTCTGCTGACTGTGAGACGAAACCTTCAATGGGCTG 1439

QY 1407 GGACGTACTGTGTGAACCTCACCTGGGGATGACAAAGCCTGGCTCTCACGAGCACCC 1466

Db 1440 GGACGTACTGTGTGAACCTCACCTGGGGATGACAAAGCCTGGCTCTCACGAGCACCC 1499

QY 1467 TGATTTCTGTTCTGACAGAGACCCAGCCTCGCCTTAAAGGATGGCAACAGTGCCTGA 1526

Db 1500 TGATTTCTGTTCTGACAGAGACCCAGCCTCGCCTTAAAGGATGGCAACAGTGCCTGA 1559

QY 1527 TCTCCGTTGGCTGCTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTAACAAAAAC 1586

Db 1560 TCTCCGTTGGCTGCTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTAACAAAAAC 1619

QY 1587 ACAAGGAATACAACCCAAATAGAAAATAGTCTCTGGGAATGTGGTCAGAAAGCAAGGCCTGA 1646

Db 1620 ACAAGGAATACAACCCAAATAGAAAATAGTCTCTGGGAATGTGGTCAGAAAGCAAGGCCTGA 1679

QY 1647 GTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTCCGGGAAACCCAGGAAAGGATCCGC 1706

Db 1680 GTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTCCGGGAAACCCAGGAAAGGATCCGC 1739

QY 1707 TACTCAAAAACCAAGAATTTAAAGGAGTTTCTTAAATTTTCGACCTTGTTCCTGAAGCTCA 1766

Db 1740 TACTCAAAAACCAAGAATTTAAAGGAGTTTCTTAAATTTTCGACCTTGTTCCTGAAGCTCA 1799

QY 1767 CTTTTCAGTGCCATTGATGTGAGATGTCTGAGTGGCTATTAAACCTTTTTTTCCTAAAG 1826

Db 1800 CTTTTCAGTGCCATTGATGTGAGATGTCTGAGTGGCTATTAAACCTTTTTTTCCTAAAG 1859

QY 1827 ATTATTGTTAAATAGATATTGTGGTTTGGGGAAGTTGAATTTTTTATAGGTTAAATGTCA 1886

Db 1860 ATTATTGTTAAATAGATATTGTGGTTTGGGGAAGTTGAATTTTTTATAGGTTAAATGTCA 1919

QY 1887 TTTTAGAGATGGGAGAGGGATTATATCTGCAGCAGCTTCAGCCATGTTGTGAAACTGAT 1946

Db 1920 TTTTAGAGATGGGAGAGGGATTATATCTGCAGCAGCTTCAGCCATGTTGTGAAACTGAT 1979

QY 1947 AAAAGCAACTTAGCAAGGCTTCTTTTCATTATTTTTTATGTTTTCACCTATATAAGTCTTAG 2006

Db 1980 AAAAGCAACTTAGCAAGGCTTCTTTTCATTATTTTTTATGTTTTCACCTATATAAGTCTTAG 2039

QY 2007 GTAAC TAGTAGATAGAAACACACTGTGTCCCAGAGTAAGGAGAGAAGCTACTATTGATTA 2066

Db 2040 GTAAC TAGTAGATAGAAACACACTGTGTCCCAGAGTAAGGAGAGAAGCTACTATTGATTA 2099

QY 2067 GAGCCTAACCCAGGTTAACTGCAAGAAGAGGGGGGATACITTCAGCTTTCCATGTAAC TG 2126

Db 2100 GAGCCTAACCCAGGTTAACTGCAAGAAGAGGGGGGATACITTCAGCTTTCCATGTAAC TG 2159

QY 2127 TATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCA 2186

Db 2160 TATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCA 2219

QY 2187 CTTCAATACACACTCATGAACCTCCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGA 2246

Db 2220 CTTCAATACACACTCATGAACCTCCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGA 2279

QY 2247 TGTGCACACTTGTGTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTG 2306

Db 2280 TGTGCACACTTGTGTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGGAGTATTTG 2339

QY 2307 GTGACAAACCTACTTTTGCTTGGCTGAGTGAAGGAATGATATTCATATTTATTTCCA 2366

Db 2340 GTGACAAACCTACTTTTGCTTGGCTGAGTGAAGGAATGATATTCATATTTATTTCCA 2399

QY 2367 TGGACATTTAGTTAGTCTTTTATATATACAGGCAATGATCTGAGTGACACTCTTGTGTA 2426

Db 2400 TGGACATTTAGTTAGTCTTTTATATATACAGGCAATGATCTGAGTGACACTCTTGTGTA 2459

QY 2427 TATTTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAAATATTAAGACTTTCC 2486

Db 2460 TATTTCCAAATTTTGTATAGTCGCTGCACATATTTGAAATC-ATATATTAAGACTTTCC 2518

QY 2487 AAAAAATTTGGTCCCTGGTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTG 2546

Db 2519 AAAGATGAGGTCCCTGGTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTG 2578

QY 2547 GAACTAAAAACCATTTACTATATGTTAGACAGACATTTTTTTTTTTTCTCTCTGAAAAA 2606

Db 2579 TAACATAAAACCATCTACTATATGTTAGACATGACATTCCTTTTCTCTCTCTGAAAAA 2638

QY 2607 -AAATGAGGGAAGAGACAAAAAATAAAAAA 2636

Db 2639 TAAAGTGTGGGAAGAGACAAAAAATAAAAAA 2669

RESULT 6

ABX76321

ID ABX76321 standard; DNA; 2669 BP.

XX ABX76321;

XX

DT 02-APR-2003 (first entry)

XX

DE Lung cancer-associated polynucleotide #185.

XX

KW Lung cancer-associated polynucleotide; gene; ds; cytostatic; emphysema; antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis; small cell lung cancer; benign lesion; precancerous lesion; bronchitis; chronic obstructive pulmonary disease; hypersensitivity pneumonitis; interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.

XX Unidentified.

OS

XX WO200286443-A2.

PN

XX 31-OCT-2002.

XX

PF 18-APR-2002; 2002WO-US012476.

XX

PR 18-APR-2001; 2001US-0284770P.

PR 10-MAY-2001; 2001US-0290492P.

PR 09-NOV-2001; 2001US-0339245P.

PR	13-NOV-2001; 2001US-0350666P.
PR	29-NOV-2001; 2001US-0334370P.
PR	12-APR-2002; 2002US-0372246P.
XX	
PA	(EOSB-) EOS BIOTECHNOLOGY INC.
XX	
PI	Aziz N, Murray R;
XX	
DR	WPI; 2003-093161/08.
DR	P-PSDB: ABU56592.

Detecting a lung cancer-associated transcript in a cell from a patient for treating lung cancer, by contacting a biological sample from the patient with a polynucleotide that exhibits increased or decreased expression in lung cancer.

Claim 22: Page 327-328: 453pp: English.

The invention relates to a method for detecting a lung cancer-associated transcript in a cell from a patient, comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridises to a sequence that is at least 80 % identical to a gene that exhibits increased or decreased expression in lung cancer samples. Lung cancer-associated polynucleotides and polypeptides are used for identifying a compound that modulates a lung cancer-associated polypeptide, for inhibiting proliferation of a lung cancer-associated cell to treat lung cancer in a patient and for treating a mammal having lung cancer by administering a modulatory compound identified. The methods are useful for treating lung cancer, such as small cell lung cancer, non-small cell lung cancer or other benign or precancerous lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides and polypeptides are useful for diagnostic purposes and as targets for screening for therapeutic compounds that modulate lung cancer, such as antibodies. Sequences ABX76124-ABX76474 represent lung cancer-associated polynucleotides of the invention.

Sequence 2669 BP; 752 A; 595 C; 597 G; 725 T; 0 U; 0 Other;

Query Match	95.7%;	Score 2547.8;	DB 7;	Length 2669;
Best Local Similarity	99.2%;	Pred. NO. 0;		
Matches 2591;	Conservative	0;	Mismatches	17;
			Indels	3;
			Gaps	3;

QY	28	AACCTTGGTGGCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCTCGGAT	87
Db	60	ACCTTGAGTGCCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCTCGGAT	119
QY	88	TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAAACGATTTTCATGATGTGCTGG	147
Db	120	TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAAACGATTTTCATGATGTGCTGG	179
QY	148	GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTAAATGGCTGGTCTTCTG	207
Db	180	GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTAAATGGCTGGTCTTCTG	239
QY	208	ATGAAAAATGACTGGAATGAAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGGA	267
Db	240	ATGAAAAATGACTGGAATGAAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGGA	299
QY	268	AAAACTCCTGGAAGGAGGCCGTGTGCAGGCGGTCTCTGACCAGTGACTCACCAGCCCTCG	327
Db	300	AAAACTCCTGGAAGGAGGCCGTGTGCAGGCGGTCTCTGACCAGTGACTCACCAGCCCTCG	359
QY	328	TGGGCTCAAAATATAACATTTGCCGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG	387
Db	360	TGGGCTCAAAATATAACATTTGCCGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG	419
QY	388	CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC	447
Db	420	CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC	479
QY	448	CATATGTTTACAACCTGGACAGCATGGTTCAGAGGACAGTGCAGGGGAAAAATGGCACCGGCC	507

Db 1560 TCTCCGTTGGTGCTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAC 1619

QY 1587 ACAAGGAATACAACCCCAATAGAAAAATAGTCTGGGAATGTGGTCAGAAGCAAAAGCCCTGA 1646

Db 1620 ACAAGGAATACAACCCCAATAGAAAAATAGTCTGGGAATGTGGTCAGAAGCAAAAGCCCTGA 1679

QY 1647 GTGTCCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGAAACCCAGGAAAAGGATCCGC 1706

Db 1680 GTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGAAACCCAGGAAAAGGATCCGC 1739

QY 1707 TACTCAAAAAACCAAGAAATTTAAAGGAGTTTCTTAAATTCGACCTGTGTTCTGAAGCTCA 1766

Db 1740 TACTCAAAAAACCAAGAAATTTAAAGGAGTTTCTTAAATTCGACCTGTGTTCTGAAGCTCA 1799

QY 1767 CTTTTAGTGCCTATTGATGTGAGATGTGCTGGAGTGGCTATTAAACCTTTTTTTCCTAAAG 1826

Db 1800 CTTTTAGTGCCTATTGATGTGAGATGTGCTGGAGTGGCTATTAAACCTTTTTTTCCTAAAG 1859

QY 1827 ATTATTGTTAAATAGATATTGTGGTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA 1886

Db 1860 ATTATTGTTAAATAGATATTGTGGTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA 1919

QY 1887 TTTTAGAGATGGGAGAGGGATTATACTGCAGGCAGCTTCAGCCATGTTGTGAAAACCTGAT 1946

Db 1920 TTTTAGAGATGGGAGAGGGATTATACTGCAGGCAGCTTCAGCCATGTTGTGAAAACCTGAT 1979

QY 1947 AAAAGCAACTTAGCAAGGCTTCTTTTCATTATTTTATGTTTCACTTATAAAGTCTTAG 2006

Db 1980 AAAAGCAACTTAGCAAGGCTTCTTTTCATTATTTTATGTTTCACTTATAAAGTCTTAG 2039

QY 2007 GTAACCTAGTAGGATAGAAACACTGTGTCCCAGAGATGAGAGAGAAGCTACTATTGATTA 2066

Db 2040 GTAACCTAGTAGGATAGAAACACTGTGTCCCAGAGATGAGAGAGAAGCTACTATTGATTA 2099

QY 2067 GAGCCTAACCCAGGTTAACTGCAAGAAGAGCGGGATACCTTTCAGCTTTCATGTAACTG 2126

Db 2100 GAGCCTAACCCAGGTTAACTGCAAGAAGAGCGGGATACCTTTCAGCTTTCATGTAACTG 2159

QY 2127 TATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCA 2186

Db 2160 TATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCA 2219

QY 2187 CTTCAATACACACTCATGAACCTCCTGTATGGAACAATAACAGGCCCAAGCCTGTGATGA 2246

Db 2220 CTTCAATACACACTCATGAACCTCCTGTATGGAACAATAACAGGCCCAAGCCTGTGATGA 2279

QY 2247 TGTGCACACTTGCTAGACTCAGAAAAAATACTACTCTCATATAAATGGTGGGAGTATTTG 2306

Db 2280 TGTGCACACTTGCTAGACTCAGAAAAAATACTACTCTCATATAAATGGTGGGAGTATTTG 2339

QY 2307 GTGACAAACCTACTTTGCTGGCTGAGTGAAGGAATGATATTCATATATTCAATTATCCCA 2366

Db 2340 GTGACAAACCTACTTTGCTGGCTGAGTGAAGGAATGATATTCATATATTCAATTATCCCA 2399

QY 2367 TGGACATTTAGTTAGTCTTTTATATATACCAGGCATGATGCTGAGTGACACTCTTGTTGA 2426

Db 2400 TGGACATTTAGTTAGTCTTTTATATATACCAGGCATGATGCTGAGTGACACTCTTGTTGA 2459

QY 2427 TATTTCCAAATTTTGPATAGTCGCTGCACATATTTGAAATCAAAATATTAAGACTTTCC 2486

Db 2460 TATTTCCAAATTTTGPATAGTCGCTGCACATATTTGAAATC-ATATATTAAAGACTTTCC 2518

QY 2487 AAAAATTTGGTCCCTGGTTTTTTCATGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTG 2546

Db 2519 AAAGATGAGGTCCTCGTTTTCATGGCAACTTGATCAGTAAGGATTTCACTCTGTTTG 2578

QY 2547 GAACTAAAAACCATTTACTATATGTTAGACAAGACATTTTTTTTTTTTTCCTCCTGAAAAA 2606

Db 2579 TAACTAAAAACCATTTACTATATGTTAGACATGACATTTCTTTTCTCTCCTCCTGAAAAA 2638

QY 2607 -AAATGAGGGAAGAGACAAAAA 2636

Db 2639 TAAAGTGTGGAGAGACAAAAA 2669

RESULT 7

ADD78274

ID ADD78274 standard; DNA; 2728 BP.

XX

AC ADD78274;

XX

DT 29-JAN-2004 (first entry)

XX

DE Human CGDD-16 coding sequence.

XX

KW Anabolic; Hypertensive; Respiratory; Anti-HIV; Antiallergic;

KW Neuroprotective; Nootropic; Antianemic; Antiarteriosclerotic;

KW Antiinflammatory; Ophthalmological; Muscular; Hepatototropic;

KW Neuroprotective; Antiasthmatic; Anticonvulsant; Virucide; Antibacterial;

KW Fungicide; Antiparasitic; Protozoacide; Antihelminthic; Cytostatic;

KW Cerebroprotective; Antiparkinsonian; Antipsoriatic; Antigout;

KW Antidiabetic; Antiarthritic; Antirheumatic; Osteopathic; Gene therapy;

KW human; cell growth; cell differentiation; cell death; CGDD;

KW cell proliferative disorder; cancer; developmental disorder;

KW neurological disorder; autoimmune disorder; inflammatory disorder;

KW infection; reproductive disorder; gene; ds.

XX

OS Homo sapiens.

XX

PN WO2003077875-A2.

XX

PD 25-SEP-2003.

XX

PF 14-MAR-2003; 2003WO-US008310.

XX

PR 15-MAR-2002; 2002US-0364494P.

PR 29-MAR-2002; 2002US-0369129P.

PR 12-APR-2002; 2002US-0372511P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

PI Kable AE, Tran UK, Hafalia AJA, Burford N, Honchell CD;

PI Lehr-Mason PM, Duggan BM, Ramkumar J, Griffin JA, Richardson TW;

PI Elliott VS, Jiang X, Jackson AA, Marquis JP, Chawla NK, Khare R;

PI Becha SD, Lee SY, Swarnakar A, Yue H, Warren BA, Baughn MR, Lal PG;

XX

Lee S, Ho A, Gandhi AR, Yao MG;

DR WPI; 2003-779081/73.

XX

P-PSDB; ADD78235.

PT New polypeptides and polynucleotides associated with cell growth,

PT differentiation and death, useful for diagnosing, treating or preventing

PT e.g. developmental, neurological, autoimmune, inflammatory or

XX reproductive disorders.

PS Claim 5; SEQ ID NO 55; 320pp; English.

XX

CC The present invention relates to novel human proteins (I; ADD78220-

CC ADD78258) and their coding sequences (II; ADD78259-ADD78297), which are

CC associated with cell growth, differentiation and death, referred to as

CC CGDD-n proteins, where n is a number from 1 to 39. The CGDD proteins and

CC their coding sequences are useful for diagnosing, treating or preventing

CC cell proliferative disorders (e.g. cirrhosis, hepatitis,

CC arteriosclerosis, psoriasis, primary thrombocytopenia) or cancers (e.g.

CC adenocarcinoma, sarcoma or cancers of the bone, bone marrow, brain,

CC breast, colon, kidney, liver, lung or uterus), developmental disorders

CC (e.g. renal tubular acidosis, Becker muscular dystrophy, gonadal

CC dysgenesis, hypothyroidism or seizures), neurological disorders (e.g.

CC Pick's disease, cataract, epilepsy, ischemic cerebrovascular disease,

CC stroke, Alzheimer's disease, Parkinson's disease or dementia),

CC autoimmune/inflammatory disorders (e.g. AIDS, allergies, asthma,

CC diabetes mellitus, bronchitis, osteoporosis, osteoarthritis, rheumatoid

CC arthritis, contact dermatitis or gout), viral, bacterial, fungal,

CC parasitic, protozoan or helminthic infections, reproductive disorders

CC (e.g. infertility, ectopic pregnancy, premature ovarian failure, delayed

CC puberty or prostatitis) or disorders of the placenta (e.g. preeclampsia,

CC choriocarcinoma, placenta previa, placental or maternal floor infarction
CC or chronic villitis).

Sequence 2728 BP; 778 A; 610 C; 607 G; 733 T; 0 U; 0 Other;

Query Match 94.2%; Score 2505.6; DB 9; Length 2728;

Best Local Similarity 97.8%; Pred. No. 0;

Matches 2596;	Conservative	0;	Mismatches	19;	Indels	39;	Gaps	4;
---------------	--------------	----	------------	-----	--------	-----	------	----

28 AACCTTGGTGCCCTGGTCCGTGAGAAATCAGCATGGAATGCTCTACTATTTCCTGGGAT 87

Db
76 ACCTTGAGTGCCCTGCCGTCCGTGAGAAATTCAGCATGGAATGCTCTACTATTTCTCTGGGAT 13

QY
88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGCCACTTGATGCCGCCAAACGATTTCATGATGTGCTGG 14

Db 136 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAACGATTTTCATGATGTGCTGG 19

[illegible]

QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGAGCACAATCAATTAAATGGCTGGTCTTCG 20

[illegible]

Db 196 GCAATGAAAGACCTTCTGCTTACATGAGGAGCACAAATCAATTAATGGCTGGTCTCTG 23

...
308 AECATACCTCATCTATAAATTGAGCGGGAGACATGAGGTGA 26

QY 208 ATGAAATGAC TGGAA TGAATAAC TAC CAG TGTGAAAGCGGGAACATGAGGTGCA 209

356 ATCAAAATGACTGGAATCAAAAACCTCTACCCAGTGTGGAGCGGGGAGACATGAGGTGGA 31

DD
250 AIGATTAAGACGGGAAAGAATTCCTACTCCCACTGTGCCTTGAATAAATTGTTTGCGCCTT

268 AAAA C T C C T G G A A G G G A G G C C G T G T G C A G G C G G T C C T G A C C A G T G A C T C A C C A G C C C T C G 32

[illegible]

316 AAAACTCTGGAAGGAGGCCGTGTGCAGGCGTCTGACCACTCAGCCCTCG 37

0

328 TGGGCTCAAATATAACATTTGCGGTGAACCTGTATATCCCTAGATGCCCAAAGGAAGATG 388

Db 376 TGGGCTCAAATATAACATTTGCCGGTGAACCTGATATTCCCTAGATGCCAAAGGAAGATG 433

QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 444

Db 436 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 455

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Qy 448 CATATGTTTACAACCTGGACAGCATGGTCAGAGGACAGTCAGCGGAAAATGGCACCGGCC 500

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

Db 496 CGTATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGACGGGAAATGGCACCGGC

2000
 2001
 2002
 2003
 2004
 2005
 2006
 2007
 2008
 2009
 2010
 2011
 2012
 2013
 2014
 2015
 2016
 2017
 2018
 2019
 2020
 2021
 2022
 2023
 2024
 2025
 2026
 2027
 2028
 2029
 2030
 2031
 2032
 2033
 2034
 2035
 2036
 2037
 2038
 2039
 2040
 2041
 2042
 2043
 2044
 2045
 2046
 2047
 2048
 2049
 2050
 2051
 2052
 2053
 2054
 2055
 2056
 2057
 2058
 2059
 2060
 2061
 2062
 2063
 2064
 2065
 2066
 2067
 2068
 2069
 2070
 2071
 2072
 2073
 2074
 2075
 2076
 2077
 2078
 2079
 2080
 2081
 2082
 2083
 2084
 2085
 2086
 2087
 2088
 2089
 2090
 2091
 2092
 2093
 2094
 2095
 2096
 2097
 2098
 2099
 2100
 2101
 2102
 2103
 2104
 2105
 2106
 2107
 2108
 2109
 2110
 2111
 2112
 2113
 2114
 2115
 2116
 2117
 2118
 2119
 2120
 2121
 2122
 2123
 2124
 2125
 2126
 2127
 2128
 2129
 2130
 2131
 2132
 2133
 2134
 2135
 2136
 2137
 2138
 2139
 2140
 2141
 2142
 2143
 2144
 2145
 2146
 2147
 2148
 2149
 2150
 2151
 2152
 2153
 2154
 2155
 2156
 2157
 2158
 2159
 2160
 2161
 2162
 2163
 2164
 2165
 2166
 2167
 2168
 2169
 2170
 2171
 2172
 2173
 2174
 2175
 2176
 2177
 2178
 2179
 2180
 2181
 2182
 2183
 2184
 2185
 2186
 2187
 2188
 2189
 2190
 2191
 2192
 2193
 2194
 2195
 2196
 2197
 2198
 2199
 2200
 2201
 2202
 2203
 2204
 2205
 2206
 2207
 2208
 2209
 2210
 2211
 2212
 2213
 2214
 2215
 2216
 2217
 2218
 2219
 2220
 2221
 2222
 2223
 2224
 2225
 2226
 2227
 2228
 2229
 2230
 2231
 2232
 2233
 2234
 2235
 2236
 2237
 2238
 2239
 2240
 2241
 2242
 2243
 2244
 2245
 2246
 2247
 2248
 2249
 2250
 2251
 2252
 2253
 2254
 2255
 2256
 2257
 2258
 2259
 2260
 2261
 2262
 2263
 2264
 2265
 2266
 2267
 2268
 2269
 2270
 2271
 2272
 2273
 2274
 2275
 2276
 2277
 2278
 2279
 2280
 2281
 2282
 2283
 2284
 2285
 2286
 2287
 2288
 2289
 2290
 2291
 2292
 2293
 2294
 2295
 2296
 2297
 2298
 2299
 2300
 2301
 2302
 2303
 2304
 2305
 2306
 2307
 2308
 2309
 2310
 2311
 2312
 2313
 2314
 2315
 2316
 2317
 2318
 2319
 2320
 2321
 2322
 2323
 2324
 2325
 2326
 2327
 2328
 2329
 2330
 2331
 2332
 2333
 2334
 2335
 2336
 2337
 2338
 2339
 2340
 2341
 2342
 2343
 2344
 2345
 2346
 2347
 2348
 2349
 2350
 2351
 2352
 2353
 2354
 2355
 2356
 2357
 2358
 2359
 2360
 2361
 2362
 2363
 2364
 2365
 2366
 2367
 2368
 2369
 2370
 2371
 2372
 2373
 2374
 2375
 2376
 2377
 2378
 2379
 2380
 2381
 2382
 2383
 2384
 2385
 2386
 2387
 2388
 2389
 2390
 2391
 2392
 2393
 2394
 2395
 2396
 2397
 2398
 2399
 2400
 2401
 2402
 2403
 2404
 2405
 2406
 2407
 2408
 2409
 2410
 2411
 2412
 2413
 2414
 2415
 2416
 2417
 2418
 2419
 2420
 2421
 2422
 2423
 2424
 2425
 2426
 2427
 2428
 2429
 2430
 2431
 2432
 2433
 2434
 2435
 2436
 2437
 2438
 2439
 2440
 2441
 2442
 2443
 2444
 2445
 2446
 2447
 2448
 2449
 2450
 2451
 2452
 2453
 2454

Qy 508 AAAGCCATCATAAACGTCCTCCCTGATGGGAACCTTTCTCACCACCCCGATGGAGAA 509

QY 268 AAAACTCCTGGAAGGAGGCCGTGTGAGGCGGTCTCTGACCGCGGTCTCTGACCGAGTGAATCACCAGCCCTCG 327
Db |||||
290 AAAACTCCTGGAAGGAGGCCGTGTGAGGCGGTCTCTGACCGAGTGAATCACCAGCCCTCG 349
QY 328 TGGGCTCAAATATAAATTTGCGGTGAACCTGTATATCCCTAGATGCAAAAGGAAGATG 387
Db |||||
350 TGGGCTCAAATATAAATTTGCGGTGAACCTGTATATCCCTAGATGCAAAAGGAAGATG 409
QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 447
Db |||||
410 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 469
QY 448 CATATGTTTACAACTGGACAGCATGGTCCAGAGACAGTACGGGGGAAATGGCACCGGCC 507
Db |||||
470 CGTATGTTTACAACTGGACAGCATGGTCCAGAGACAGTACGGGGGAAATGGCACCGGCC 529
QY 508 AAAGCCATCATACGCTCTCCCTGATGGGAAACCTTTTCTCACCCACCCCGGATGGAGAA 567
Db |||||
530 AAAGCCATCATACGCTCTCCCTGATGGGAAACCTTTTCTCACCCACCCCGGATGGAGAA 589
QY 568 GATGGAATTTCACTAGCTCTTCCACACACTTGGTCAAGTATTTCCAGAAATGGGACGAT 627
Db |||||
590 GATGGAATTTCACTAGCTCTTCCACACACTTGGTCAAGTATTTCCAGAAATGGGACGAT 649
QY 628 GTTCAGTGAGAGTTTCTGTGAACACACAGCCAAATGTGACACTTGGGCCTCAACTCATGGAAG 687
Db |||||
650 GTTCAGTGAGAGTTTCTGTGAACACACAGCCAAATGTGACACTTGGGCCTCAACTCATGGAAG 709
QY 688 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 747
Db |||||
710 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 769
QY 748 ACGTGGTAACAGATCAGATTCCTGTGTTTGTGACTATGTTCCAGAAAGACGATCGAAAT 807
Db |||||
770 ACGTGGTAACAGATCAGATTCCTGTGTTTGTGACTATGTTCCAGAAAGACGATCGAAAT 829
QY 808 CATCCGACGAAACCTTCC - CAAAGATCTCCCCATATGTTTGATGTCCTGATTCATGATC 866
Db |||||
830 CATCCGACGAAACCTTCTCAAAGATCTCCCCATATGTTTGATGTCCTGATTCATGATC 889
QY 867 CTAGCCACTTCCTCAATTAATCTACCATTAACATAAGTGGAGCTTCGGGGATAAATACTG 926
Db |||||
890 CTAGCCACTTCCTCAATTAATCTACCATTAACATAAGTGGAGCTTCGGGGATAAATACTG 949
QY 927 GCCTGTTGTTTCCACCAATCATACTGTGAATCAGACGTATGTGCTCAATGGAACCTTCA 986
Db |||||
950 GCCTGTTGTTTCCACCAATCATACTGTGAATCAGACGTATGTGCTCAATGGAACCTTCA 1009
QY 987 GCCTTAAACCTCACTGTGAAAGCTGCAGCACCGACCTGTCCGCCACCGCCACCAAC 1046
Db |||||
1010 GCCTTAAACCTCACTGTGAAAGCTGCAGCACCGACCTGTCCGCCACCGCCACCAAC 1069
QY 1047 CCAGACCTTCAAAACCCCTTCTTT - - - - - 1074
Db |||||
1070 CCAGACCTTCAAAACCCCTTCTTTTAGCAACTACTCTAAATCTTATGATTCAAACA 1129
QY 1075 - - - - - AGGACCTGCTGTGACAAACCCCTTCTTTTAGCAACTACTCTAAATCTTATGATTCAAACA 1130
Db |||||
1130 CCCCAGGACCTACTGGTGACAAACCCCTTCTTTAGCAACTACTCTAAATCTTATGATTCAAACA 1189
QY 1131 AGATTAAACAGATATGGCCACTTTCAAGCCACCATCAATTTGTAGAGGAACTTTAGAGG 1190
Db |||||
1190 AGATTAAACAGATATGGCCACTTTCAAGCCACCATCAATTTGTAGAGGAACTTTAGAGG 1249
QY 1191 TTAACATCATCCAGATGACAGACGTCTGATCCCGTGCCATGGCCTGAAAGCTCCCTAA 1250
Db |||||
1250 TTAACATCATCCAGATGACAGACGTCTGATCCCGTGCCATGGCCTGAAAGCTCCCTAA 1309
QY 1251 TAGACTTTGTGTCGACCTGCCAAGGAGCATTCCTCCACGGAGTCTGTACCATCATTTCTG 1310
Db |||||
1310 TAGACTTTGTGTCGACCTGCCAAGGAGCATTCCTCCACGGAGTCTGTACCATCATTTCTG 1369

QY 1311 ACCCACCTGCGAGATCACCCAGAAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT 1370
Db |||||
1370 ACCCACCTGCGAGATCACCCAGAAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT 1429
QY 1371 GTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGACGTACTGTGTGAACCTCACCC 1430
Db |||||
1430 GTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGACGTACTGTGTGAACCTCACCC 1489
QY 1431 TGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCTCTGACAGAC 1490
Db |||||
1490 TGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCTCTGACAGAC 1549
QY 1491 CAGCCTCGCTTTAAGGATGGCAACACAGTCCCTGATCTCCGTTGGCTGTGGCCATAT 1550
Db |||||
1550 CAGCCTCGCTTTAAGGATGGCAACACAGTCCCTGATCTCCGTTGGCTGTGGCCATAT 1609
QY 1551 TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAAGGAATACAAACCAATAGAAA 1610
Db |||||
1610 TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAAGGAATACAAACCAATAGAAA 1669
QY 1611 ATAGTCTGGGAATGGTTCAGAAAGCAAGGCTGAGTGTCTTTCTCAACCGTGCAAAAG 1670
Db |||||
1670 ATAGTCTGGGAATGGTTCAGAAAGCAAGGCTGAGTGTCTTTCTCAACCGTGCAAAAG 1729
QY 1671 CCGTGTCTTCCCGGAAAACAGGAAAGATCCGCTACTCAAAAAACCAAGAAATTTAAAG 1730
Db |||||
1730 CCGTGTCTTCCCGGAAAACAGGAAAGATCCGCTACTCAAAAAACCAAGAAATTTAAAG 1789
QY 1731 GAGTTTCTTAAATTCGACCTTGTTCGAAAGCTCACTTTTTCAGTGCCATGTGAGA 1790
Db |||||
1790 GAGTTTCTTAAATTCGACCTTGTTCGAAAGCTCACTTTTTCAGTGCCATGTGAGA 1849
QY 1791 TGTGCTGGAGTGCCTATTAAACCTTTTCTTAAAGATTTATGTTAAATAGATATTGTGG 1850
Db |||||
1850 TGTGCTGGAGTGCCTATTAAACCTTTTCTTAAAGATTTATGTTAAATAGATATTGTGG 1909
QY 1851 TTTGGGGAAGTTGAATTTTTTATAGGTTAAATGTCACTTTTAGAGATGGGGAGGGATTA 1910
Db |||||
1910 TTTGGGGAAGTTGAATTTTTTATAGGTTAAATGTCACTTTTAGAGATGGGGAGGGATTA 1969
QY 1911 TACTGCAGGCAGCTTCAGCCATGTTGTGAAAGCTGATAAAGCAACTTAGCAAGGCTTCTT 1970
Db |||||
1970 TACTGCAGGCAGCTTCAGCCATGTTGTGAAAGCTGATAAAGCAACTTAGCAAGGCTTCTT 2029
QY 1971 TTCATTATTTTTTATGTTTCACTTATAAAGCTTAGGTAACTAGTAGGATAGAAAACACTG 2030
Db |||||
2030 TTCATTATTTTTTATGTTTCACTTATAAAGCTTAGGTAACTAGTAGGATAGAAAACACTG 2089
QY 2031 TGTCCCGAGAGTAAGGAGAGAAGCTACTATTGATAGAGCCCTAACCCAGGTTAACTGCAA 2090
Db |||||
2090 TGTCCCGAGAGTAAGGAGAGAAGCTACTATTGATAGAGCCCTAACCCAGGTTAACTGCAA 2149
QY 2091 GAAGAGCGGGATACCTTTCAGCTTTCCATGTAAGTGTATGATGATAAAGCCAACTGAGTCCA 2150
Db |||||
2150 GAAGAGCGGGATACCTTTCAGCTTTCCATGTAAGTGTATGATAAAGCCAACTGAGTCCA 2209
QY 2151 GTTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTCC 2210
Db |||||
2210 GTTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTCC 2269
QY 2211 TGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGTCTAGACTCAGAA 2270
Db |||||
2270 TGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGTCTAGACTCAGAA 2329
QY 2271 AAAATACTACTCTCATAAATGGGTGGGAGTATTTTGGTGACAACTACTTTTGGCTG 2330
Db |||||
2330 AAAATACTACTCTCATAAATGGGTGGGAGTATTTTGGTGACAACTACTTTTGGCTG 2389
QY 2331 AGTGAAGGAATGATATTCATATATTTCACTTTATTCATGGACATTTAGTTAGTCTTTT 2390
Db |||||
2390 AGTGAAGGAATGATATTCATATATTTCACTTTATTCATGGACATTTAGTTAGTCTTTT 2449
QY 2391 TATACAGGCATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTATAGTGC 2450

Db 2450 TATACCAGGCATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTACAGTCG 2509
QY 2451 CTGCACATATTTGAAATCAAATATTAAGACTTTCCAAAAATTTGGTCCTGGTTTTTCA 2510
Db 2510 CTGCACATATTTGAAATC-ATATATTAAGACTTTCCAAAGATGAGGTCCTGGTTTTTCA 2568
QY 2511 TGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTGGAACTAAAAACCAATTTACTATATGT 2570
Db 2569 TGGCAACTTGATCAGTAAGGATTTCACTCTGTTTGTAACTAAAAACCATCTACTATATGT 2628
QY 2571 TAGACAAGACATTTTTTTTTTTTTCCTTCCCTGAAAAA-AAAAATGAGGGAAGAGACA 2624
Db 2629 TAGACATGACATTTCTTTTCTCTCCTTCCCTGAAAAATAAAGTGTGGGAAGAGACA 2683

RESULT 9
ACA66861
ID ACA66861 standard; cDNA; 2683 BP.
XX
AC ACA66861;
XX
DT 23-JUN-2003 (first entry)
XX
DE cDNA encoding human PRO polypeptide #21.
DE
KW Human; PRO polypeptide; secreted and transmembrane protein;
KW anti-PRO antibody; diagnostic assay; gene expression; tumour; cytostatic;
KW gene; ss.
KW
OS Homo sapiens.
XX
PN US2003036635-A1.
XX
PD 20-FEB-2003.
XX
PF 28-AUG-2002; 2002US-00230163.
XX
PR 25-JUL-2000; 2000US-0220638P.
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX

WPI; 2003-342045/32.
DR P-PSDB; ABU80759.
XX

PT One hundred and twenty two nucleic acids encoding PRO polypeptides,
PT useful for the manufacture of a medicament for diagnosing or treating
PT tumor.
XX

PS Claim 2; Fig 41; 314pp; English.
XX

CC The present invention relates to the isolation of novel human PRO
CC polypeptides, and the polynucleotide sequences encoding them. The PRO
CC polypeptides are secreted and transmembrane proteins. The PRO
CC polypeptides and polynucleotides are useful for preparing a medicament
CC useful in the diagnosis and treatment of tumours. Anti-PRO antibodies are
CC useful in diagnostic assays for PRO, by detecting its expression in
CC specific cells, tissues or serum, and for affinity purification of PRO
CC from recombinant cell culture or natural sources. ACA66841-ACA66962
CC represent cDNA sequences encoding the human PRO polypeptides of the
CC invention. Note: The sequence data for this patent was obtained in
CC electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/psipsDIEntry.html
XX

SQ Sequence 2683 BP; 751 A; 604 C; 596 G; 732 T; 0 U; 0 Other;

Query Match 93.4%; Score 2485; DB 7; Length 2683;

Best Local Similarity 97.8%; Pred. No. 0;
Matches 2576; Conservative 0; Mismatches 20; Indels 39; Gaps 4;
QY 28 AACCTTGGTGCCTCGCTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCCCTGGGAT 87
Db 50 ACCTTGAGTGCCTCGCTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCCCTGGGAT 109
QY 88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATCCGCCAAACGATTTTCATGATGTCTGG 147
Db 110 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATCCGCCAAACGATTTTCATGATGTCTGG 169
QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGGTCTTCTG 207
Db 170 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGGTCTTCTG 229
QY 208 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGGA 267
Db 230 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGGA 289
QY 268 AAAACTCCTGGAAGGGAGGCCGTGTGCAGGGCGGTCTGACCATGACTCACCAGCCCTCG 327
Db 290 AAAACTCCTGGAAGGGAGGCCGTGTGCAGGGCGGTCTGACCATGACTCACCAGCCCTCG 349
QY 328 TGGGCTCAAATATAACATTTGCGGTGAACCTGATATTTCCCTAGATGCCAAAGGAAGATG 387
Db 350 TGGGCTCAAATATAACATTTGCGGTGAACCTGATATTTCCCTAGATGCCAAAGGAAGATG 409
QY 388 CCAATGGCAACATAGTCTATGAGAAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 447
Db 410 CCAATGGCAACATAGTCTATGAGAAAGAACTGCAGAAATGAGGCTGGTTTATCTGCTGATC 469
QY 448 CATATGTTTACAACCTGGACAGCATGGTGCAGAGGACAGTGCAGGGGAAATGGCACCGGCC 507
Db 470 CGTATGTTTACAACCTGGACAGCATGGTGCAGAGGACAGTGCAGGGGAAATGGCACCGGCC 529
QY 508 AAAGCCCATCATAACGCTCTCCCTGATGGGAAACCTTTTCTCTCACCACCCCGGATGGAGAA 567
Db 530 AAAGCCCATCATAACGCTCTCCCTGATGGGAAACCTTTTCTCTCACCACCCCGGATGGAGAA 589
QY 568 GATGGAATTTTCATCTACGTCTTCCACACACTTTGGTCAGTATTTCCAGAAATGGGACGAT 627
Db 590 GATGGAATTTTCATCTACGTCTTCCACACACTTTGGTCAGTATTTCCAGAAATGGGACGAT 649
QY 628 GTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCCCTCAACTCATGGAAG 687
Db 650 GTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCCCTCAACTCATGGAAG 709
QY 688 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 747
Db 710 TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 769
QY 748 ACGTGGTAACAGATCAGATTCCTGTGTTTGTGACTATGTTCCAGAAAGACGATCGAAATT 807
Db 770 ACGTGGTAACAGATCAGATTCCTGTGTTTGTGACTATGTTCCAGAAAGACGATCGAAATT 829
QY 808 CATCCGACGAAACCTTCC-CAAAGATCTCCCATTTATGTTTGTATGTCCTGATTCATGATC 866
Db 830 CATCCGACGAAACCTTCTCAAAGATCTCCCATTTATGTTTGTATGTCCTGATTCATGATC 889
QY 867 CTAGCCACTTCTCAATTTATTCTACCAATTAACACTACAAGTGGAGCTTCGGGGATAATACTG 926
Db 890 CTAGCCACTTCTCAATTTATTCTACCAATTAACACTACAAGTGGAGCTTCGGGGATAATACTG 949
QY 927 GCCTGTTTGTTCACCAATCATACTGTGAATCACACGATGTGCTCAATGGAACCTTCA 986
Db 950 GCCTGTTTGTTCACCAATCATACTGTGAATCACACGATGTGCTCAATGGAACCTTCA 1009
QY 987 GCCTTAACCTCACTGTGAAAGCTGCAGACCAGGACCTTGTCCGCCACCGCCACCACCAC 1046
Db 1010 GCCTTAACCTCACTGTGAAAGCTGCAGACCAGGACCTTGTCCGCCACCGCCACCACCAC 1069
QY 1047 CCAGACCTTCAAAACCCACCCCTTCTTT----- 1074

Db	1070	CCAGACCTTCAAAACCCACCCCTTCTTTAGCAACTACTCTAAAAATCTTATGATTCAAAACA	1129
QY	1075	---AGGACCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACTGCC	1130
Db	1130	CCCCAGGACCTACTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACTGCC	1189
QY	1131	AGATTAAACAGATATGGCCACTTTTCAAGCCACCATCACAATTTAGAGGGAATCTTAGAGG	1190
Db	1190	AGATTAAACAGATATGGCCACTTTCAAGCCACCATCACAATTTAGAGGGAATCTTAGAGG	1249
QY	1191	TTAACATATCCAGATGACAGACGTCCTGATGCCGTGCCATGGCCTGAAAGCTCCCTAA	1250
Db	1250	TTAACATATCCAGATGACAGACGTCCTGATGCCGTGCCATGGCCTGAAAGCTCCCTAA	1309
QY	1251	TAGACTTTGTGCTGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACCATCATTTCTG	1310
Db	1310	TAGACTTTGTGCTGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACCATCATTTCTG	1369
QY	1311	ACCCACCTGGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT	1370
Db	1370	ACCCACCTGGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT	1429
QY	1371	GTCTGCTGACTGTGAGACGAACCTTCAATGGTCTGGGACGTACTGTGTAAACCTCACCC	1430
Db	1430	GTCTGCTGACTGTGAGACGAACCTTCAATGGTCTGGGACGTACTGTGTAAACCTCACCC	1489
QY	1431	TGGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTCTCTGACAGAGACC	1490
Db	1490	TGGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTCTCTGACAGAGACC	1549
QY	1491	CAGCCTCGCCTTTAAGGATGGCAAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATAT	1550
Db	1550	CAGCCTCGCCTTTAAGGATGGCAAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATAT	1609
QY	1551	TTGTCACTGTGATCTCCCTTTGGTGTAACAAAAACAAGGAATACAAACCAATAGAAA	1610
Db	1610	TTGTCACTGTGATCTCCCTTTGGTGTAACAAAAACAAGGAATACAAACCAATAGAAA	1669
QY	1611	ATAGTCCTGGGAATGTGTCAGAAAGCAAAGGCCTGAGTGCTTTCTCAACCGTGCAAAAG	1670
Db	1670	ATAGTCCTGGGAATGTGTCAGAAAGCAAAGGCCTGAGTGCTTTCTCAACCGTGCAAAAG	1729
QY	1671	CCGTGTTCTTCCCGGAAACCCAGGAAAGGATCCGCTACTCAAAAACCAAGAATTTAAAG	1730
Db	1730	CCGTGTTCTTCCCGGAAACCCAGGAAAGGATCCGCTACTCAAAAACCAAGAATTTAAAG	1789
QY	1731	GAGTTTCTTAAATTTTCGACCTTGTTTCTGAAGCTCACTTTTCAGTGCCATTGATGTGAGA	1790
Db	1790	GAGTTTCTTAAATTTTCGACCTTGTTTCTGAAGCTCACTTTTCAGTGCCATTGATGTGAGA	1849
QY	1791	TGTGCTGGAGTGGCTATTAAACCTTTTTTCTTAAAGATTATTGTTAAATAGATATTGTGG	1850
Db	1850	TGTGCTGGAGTGGCTATTAAACCTTTTTTCTTAAAGATTATTGTTAAATAGATATTGTGG	1909
QY	1851	TTTGGGAAAGTTGAATTTTTTATAGGTTAAATGTCAATTTAGAGATGGGAGAGGGATTA	1910
Db	1910	TTTGGGAAAGTTGAATTTTTTATAGGTTAAATGTCAATTTAGAGATGGGAGAGGGATTA	1969
QY	1911	TACTGCAGGCAGCTTCAGCCATGTTGTGAAACTGATAAAAGCAACTTAGCAAGGCTTCTT	1970
Db	1970	TACTGCAGGCAGCTTCAGCCATGTTGTGAAACTGATAAAAGCAACTTAGCAAGGCTTCTT	2029
QY	1971	TTTCATTATTTTATGTTTTCACCTTATAAAGTCTTAGGTAAGTACTAGTAGGATAGAAACACTG	2030
Db	2030	TTTCATTATTTTATGTTTTCACCTTATAAAGTCTTAGGTAAGTACTAGTAGGATAGAAACACTG	2089
QY	2031	TGTCCCGAGAGTAAGGAGAGAAGCTACTATTGATTAGAGCCTTAACCCAGGTTAACTGCAA	2090
Db	2090	TGTCCCGAGAGTAAGGAGAGAAGCTACTATTGATTAGAGCCTTAACCCAGGTTAACTGCAA	2149
QY	2091	GAAGAGGCGGGAATACTTTCAGCTTTCCATGTAACTGTATGCATAAAGCCCAATGTAGTCCA	2150
Db	2150	GAAGAGGCGGGAATACTTTCAGCTTTCCATGTAACTGTATGCATAAAGCCCAATGTAGTCCA	2209

RESULT 10

ACD68673

ACD68613
ID ACD68613 standard: cDNA; 2683 BP.

AA ACN68613.

AC
YY
ACD0000137

DT 17-SEP-2003 (first entry)

DI 17-SEP-2003 1110C CHERI,
XX

XX	Novel human secreted and transmembrane protein PRO9925 cDNA.
DE	
XX	
KW	Human; secreted and transmembrane protein; PRO; cytostatic;
KW	antiarthritic; osteopathic; gene therapy; TNF-Agonist-Alpha;
KW	chondrocyte stimulator; pericyte stimulator; fibroblast modulator;
KW	pharmaceutical; diagnostic; biosensor; bioreactor; tumour; lung tumour;
KW	colon tumour; breast tumour; prostate tumour; rectal tumour;
KW	liver tumour; bone disorder; cartilage disorder; sports injury;
KW	arthritis; wound; gene; ss.
KW	

XXV Homo sapiens.

AA
PN
IIS2003045687-A1.

XX
PD 06-MAR-2003

XX
DE 13-AUG-2002. 2002US-00218631

XX
BB
01

PR 01-JUN-2001; 2001WO-95011800
PR 28-JUN-2001; 2001WO-95021066

PR 29-JUN-2001; 2001WC-030Z1000
PR 09-APR-2002; 2002US-00119480

XX
DZ (CETH) GENENTECH INC

xx Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
 PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
 PI xx

WPT: 2003-512315/48

DR
WEL, 2003-5125137
P-PSDB: ABO33725.

XX
XX
F-100D, AD000173.

PT

PT useful for stimulating Tumor Necrosis Factor alpha, or chondrocyte or
PT pericyte proliferation, especially for treating lung tumors, arthritis or
XX wounds in a mammal.

PS Claim 2; Fig 41; 314pp; English.

XX
CC The invention describes an isolated nucleic acid molecule comprising a
CC sequence with at least 80% identity to: (a) a nucleotide encoding any of
CC 122 PRO (secreted and transmembrane) polypeptides whose sequences are
CC fully defined in the specification; or (b) any of 122 nucleotide
CC sequences having e.g. 4834, 2504 or 1759 bp fully defined in the
CC specification; or the full length coding sequence of any these 122
CC nucleotide sequences. The PRO polypeptides or polynucleotides are useful
CC as pharmaceuticals, diagnostics, biosensors or bioreactors. These are
CC particularly useful for detecting tumours (e.g. lung tumour, colon
CC tumour, breast tumour, prostate tumour, rectal tumour, or liver tumour)
CC in a mammal, for stimulating the release of TNF-alpha from human blood,
CC for stimulating the proliferation or differentiation of chondrocyte
CC cells, for stimulating proliferation of pericyte cells, or for modulating
CC normal human dermal fibroblast proliferation. The PRO nucleic acid or
CC polypeptide is also useful for treating tumours or various bone and/or
CC cartilage disorders (e.g. sports injuries or arthritis), or wounds. The
CC PRO polypeptides are useful in drug screening, particularly as targets
CC for therapeutic intervention in these diseases, and in the diagnostic
CC determination of the presence of these diseases. The PRO polypeptides are
CC also useful as molecular weight markers, or for chromosome
CC identification. The PRO genes are useful as hybridisation probes, or for
CC screening libraries of human cDNA, genomic DNA or mRNA. The PRO genes may
CC also be used in gene therapy, particularly for replacing a defective
CC gene. This sequence encodes a novel human secreted and transmembrane PRO
CC polypeptide
XX

SQ Sequence 2683 BP; 751 A; 604 C; 596 G; 732 T; 0 U; 0 Other;

Query Match 93.4%; Score 2485; DB 7; Length 2683;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 2576; Conservative 0; Mismatches 20; Indels 39; Gaps 4;

QY	28	AACCTTGGTGCCTGCGTCCTGAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGGAT	87
Db	50	ACCTTGAGTGCCTGCGTCCTGAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGGAT	109
QY	88	TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTTCATGATGCTGG	147
Db	110	TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTTCATGATGCTGG	169
QY	148	GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTAATGGCTGCTTCTG	207
Db	170	GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAATCAATTAATGGCTGCTTCTG	229
QY	208	ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA	267
Db	230	ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA	289
QY	268	AAACTCCTGGAGGAGGCCGTGTGCAGGCGTCTGACCACTGACTCACCAGCCCTCG	327
Db	290	AAACTCCTGGAGGAGGCCGTGTGCAGGCGTCTGACCACTGACTCACCAGCCCTCG	349
QY	328	TGGGCTCAAATATAACATTTGCGGTGAACCTGATATCCCTAGATGCCAAAGGAAGATG	387
Db	350	TGGGCTCAAATATAACATTTGCGGTGAACCTGATATCCCTAGATGCCAAAGGAAGATG	409
QY	388	CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTATCTGCTGATC	447
Db	410	CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTTATCTGCTGATC	469
QY	448	CATATGTTTAACTGGACAGCATGGTCAGAGGACAGTGACGGGAAATGGCACCCGCC	507
Db	470	CGTATGTTTAACTGGACAGCATGGTCAGAGGACAGTGACGGGAAATGGCACCCGCC	529
QY	508	AAAGCCATCAACAGTCTTCCCTGATGGGAAACCTTTTCCTCACCCACCCCGGATGGAGAA	567
Db	530	AAAGCCATCAACAGTCTTCCCTGATGGGAAACCTTTTCCTCACCCACCCCGGATGGAGAA	589

QY	568	GATGGAATTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATGGGACGAT	627
Db	590	GATGGAATTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATGGGACGAT	649
QY	628	GTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG	687
Db	650	GTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG	709
QY	688	TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGACAAAGTGAAGATGTGT	747
Db	710	TGACTGTCTACAGAAGACATGGACGGGCATATGTTCCCATCGACAAAGTGAAGATGTGT	769
QY	748	ACGTGTAACAGATCAGATTCCCTGTGTTTGTGACTATGTTCCAGAAACGATCGAAAT	807
Db	770	ACGTGTAACAGATCAGATTCCCTGTGTTTGTGACTATGTTCCAGAAACGATCGAAAT	829
QY	808	CATCCGACGAAACCTTCC-CAAAGATCTCCCCATTATGTTGATGTCCTGATTCATGATC	866
Db	830	CATCCGACGAAACCTTCCCTCAAAGATCTCCCCATTATGTTGATGTCCTGATTCATGATC	889
QY	867	CTAGCCACTTCTCAATTATTCTACCAATTAACACTAAGTGGAGCTTCGGGGATAATCTG	926
Db	890	CTAGCCACTTCTCAATTATTCTACCAATTAACACTAAGTGGAGCTTCGGGGATAATCTG	949
QY	927	GCCTGTTTGTTCACCAATCATACTGTGAATCACACGATGTCTCAATGGAACCTTCA	986
Db	950	GCCTGTTTGTTCACCAATCATACTGTGAATCACACGATGTCTCAATGGAACCTTCA	1009
QY	987	GCCTTAACCTCACTGTGAAGCTGCAGCACCCAGGACCTTGTCCGCCACCCACCCAC	1046
Db	1010	GCCTTAACCTCACTGTGAAGCTGCAGCACCCAGGACCTTGTCCGCCACCCACCCAC	1069
QY	1047	CCAGACCTTCAAAACCCACCCCTTCTTT-----	1074
Db	1070	CCAGACCTTCAAAACCCACCCCTTCTTTAGCAACTACTCTAAATCTTATGATTCAAACA	1129
QY	1075	---AGGACCTGCTGTGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCC	1130
Db	1130	CCCCAGGACCTACTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCC	1189
QY	1131	AGATTAAACAGATATGGCCACTTTCAAGCCACCATCAATTTGAGAGGAACTCTTAGAGG	1190
Db	1190	AGATTAAACAGATATGGCCACTTTCAAGCCACCATCAATTTGAGAGGAACTCTTAGAGG	1249
QY	1191	TTAACATCATCCAGATGACAGACGTCTGATGCGGTGCCATGGCCTGAAAGCTCCCTAA	1250
Db	1250	TTAACATCATCCAGATGACAGACGTCTGATGCGGTGCCATGGCCTGAAAGCTCCCTAA	1309
QY	1251	TAGACTTTGTCTGACCTGCCAAGGAGCAATTCACGAGGTCTGTACCATCATTTCTG	1310
Db	1310	TAGACTTTGTCTGACCTGCCAAGGAGCAATTCACGAGGTCTGTACCATCATTTCTG	1369
QY	1311	ACCCACCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGAGATGT	1370
Db	1370	ACCCACCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGAGATGT	1429
QY	1371	GTCTGCTGACTGTGAGACGAACTTCAATGGGTCTGGACGTCGATCTGTTGGCTGCTTGGCCATAT	1430
Db	1430	GTCTGCTGACTGTGAGACGAACTTCAATGGGTCTGGACGTCGATCTGTTGGCTGCTTGGCCATAT	1489
QY	1431	TGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCTGACAGAGACC	1490
Db	1490	TGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCTGACAGAGACC	1549
QY	1491	CAGCCTCGCCTTTAAGGATGGCAACAGTCCCTGATCTCCGTTGGCTGCTTGGCCATAT	1550
Db	1550	CAGCCTCGCCTTTAAGGATGGCAACAGTCCCTGATCTCCGTTGGCTGCTTGGCCATAT	1609
QY	1551	TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAGGAATACAAACCAATAGAAA	1610
Db	1610	TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAGGAATACAAACCAATAGAAA	1669

QY	1611	ATAGTCTCTGGGAATGTGGTCAGAACGCAAAAGCCCTGAGTGTCTTCTCAACCGTGCAAAAG	1670	AC	ACA68517;
Db	1670	ATAGTCTCTGGGAATGTGGTCAGAACGCAAAAGCCCTGAGTGTCTTCTCAACCGTGCAAAAG	1729	XX	
QY	1671	CCGTGTTCTCCCGGGAACCCAGGAAAGGATCCGCTACTCAAAAACCAAGAATTTAAAG	1730	DT	25-JUN-2003 (first entry)
Db	1730	CCGTGTTCTCCCGGGAACCCAGGAAAGGATCCGCTACTCAAAAACCAAGAATTTAAAG	1789	XX	Novel human secreted and transmembrane protein PRO9925 cDNA.
QY	1731	GAGTTTCTTAATTTTCGACCTTGTTTCTGAAGCTCACTTTTCAGTGCCATTGATGTGAGA	1790	KW	Human; secreted and transmembrane protein; PRO; cardiant; cytostatic;
Db	1790	GAGTTTCTTAATTTTCGACCTTGTTTCTGAAGCTCACTTTTCAGTGCCATTGATGTGAGA	1849	KW	antiangiogenic; hypotensive; vulnenry; antiarteriosclerotic;
QY	1791	TGTGCTGGAGTGGCTATTAAACCTTTTTTCTCTAAAGATTATTGTTAAATAGATATTGTGG	1850	KW	gene therapy; cardiovascular disorder; endothelial disorder;
Db	1850	TGTGCTGGAGTGGCTATTAAACCTTTTTTCTCTAAAGATTATTGTTAAATAGATATTGTGG	1909	KW	angiogenic disorder; cardiac hypertrophy; trauma; cancer;
QY	1851	TTTGGGGAAGTTGAATTTTATAGGTTAAATGTCAATTTTAGAGATGGGAGAGGATTA	1910	KW	age-related macular degeneration; atherosclerosis; hypertension;
Db	1910	TTTGGGGAAGTTGAATTTTATAGGTTAAATGTCAATTTTAGAGATGGGAGAGGATTA	1969	KW	arterial restenosis; rheumatoid arthritis; angina; myocardial infarction;
QY	1911	TACTGCAGGCAGCTTCAGCCATGTTGTGAACTGATAAAAGCAACTTAGCAAGGCTTCTT	1970	KW	thrombophlebitis; lymphangitis; tumour angiogenesis; breast carcinoma;
Db	1970	TACTGCAGGCAGCTTCAGCCATGTTGTGAACTGATAAAAGCAACTTAGCAAGGCTTCTT	2029	KW	liver carcinoma; wound healing; chromosome mapping; gene mapping; gene;
QY	1971	TTCAATTATTTTTTATGTTTCACTTATAAAGTCTTAGGTAAGTCTAGTAGGATAGAAACACTG	2030	XX	Homo sapiens.
Db	2030	TTCAATTATTTTTTATGTTTCACTTATAAAGTCTTAGGTAAGTCTAGTAGGATAGAAACACTG	2089	OS	US2003088063-A1.
QY	2031	TGTCCCGAGAGTAAGGAGAGAAAGCTACTATTGATTAGAGCCTAACCCAGGTTAACTGCAA	2090	PN	08-MAY-2003.
Db	2090	TGTCCCGAGAGTAAGGAGAGAAAGCTACTATTGATTAGAGCCTAACCCAGGTTAACTGCAA	2149	XX	12-AUG-2002; 2002US-00219003.
QY	2091	GAAGAGCGGGGATACCTTCAGCTTTCCATGTAAGTCTAGTATGCATAAAGCCCAATGTAGTCCA	2150	PR	25-JUL-2000; 2000US-0220664P.
Db	2150	GAAGAGCGGGGATACCTTCAGCTTTCCATGTAAGTCTAGTATGCATAAAGCCCAATGTAGTCCA	2209	PR	01-JUN-2001; 2001WO-US017800.
QY	2151	GTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTCC	2210	PR	29-JUN-2001; 2001WO-US021066.
Db	2210	GTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTCC	2269	PR	09-APR-2002; 2002US-00119480.
QY	2211	TGATGGAAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGCTAGACTCAGAA	2270	XX	(GETH) GENENTECH INC.
Db	2270	TGATGGAAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGCTAGACTCAGAA	2329	PA	Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
QY	2271	AAAATACTACTCTCATATAATGGGTGGGAGTATTTTGGTGACAACTACTTTTGGCTGGCTG	2330	XX	Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
Db	2330	AAAATACTACTCTCATATAATGGGTGGGAGTATTTTGGTGACAACTACTTTTGGCTGGCTG	2389	PI	WPI; 2003-393229/37.
QY	2331	AGTGAAGGAATGATATTTCATATATTCATTTATTCATGGACATTTAGTTAGTGCCTTTTA	2390	DR	P-PSDB; ABU82068.
Db	2390	AGTGAAGGAATGATATTTCATATATTCATTTATTCATGGACATTTAGTTAGTGCCTTTTA	2449	XX	One hundred and eighty seven nucleic acids encoding PRO polypeptides,
QY	2391	TATACCAGGCATGATGCTGAGTGACACTCTTGTTGTATATTTCCAAATTTTGTATAGTCG	2450	PT	useful in diagnosis and treatment of cardiovascular (e.g. myocardial
Db	2450	TATACCAGGCATGATGCTGAGTGACACTCTTGTTGTATATTTCCAAATTTTGTATAGTCG	2509	PT	infarction), endothelial or angiogenic disorders in a mammal.
QY	2451	CTGCACATATTTGAAATCAAAATATTAAGACTTTCCAAAAATTTGGTCCCTGGTTTTTCA	2510	XX	Claim 2; Fig 41; 314pp; English.
Db	2510	CTGCACATATTTGAAATC-ATATATTAAGACTTTCCAAAGATGAGGTCCCTGGTTTTTCA	2568	XX	The invention describes one hundred and eighty seven nucleic acids
QY	2511	TGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTGGAACTAAACCATTTTACTATATGT	2570	CC	encoding novel human secreted and transmembrane (PRO) polypeptides. The
Db	2569	TGGCAACTTGATCAGTAAGGATTTTACCTCTGTTTGTAACTAAACCATCTACTATATGT	2628	CC	PRO nucleic acids, polypeptides, agonists and antagonists are useful for
QY	2571	TAGACAAGACATTTTTTTTTTCTCTCCCTGAAAAA-AAAATGAGGGGAAGAGACA	2624	CC	treating or diagnosing a cardiovascular, endothelial or angiogenic
Db	2629	TAGACATGACATCTTTTTTCTCTCTCTGAAAAATAAAGTGTGGGAAGAGACA	2683	CC	disorder in a mammal, e.g. cardiac hypertrophy, trauma, cancer, age-
RESULT 11				CC	related macular degeneration, atherosclerosis, hypertension, arterial
ACA68517				CC	restenosis, rheumatoid arthritis, angina, myocardial infarctions,
ID				CC	thrombophlebitis, lymphangitis, tumour angiogenesis (such as breast
XX				CC	carcinoma and liver carcinoma) and wound healing. The PRO nucleic acids
				CC	have applications in molecular biology, including use as hybridisation
				CC	probes, and in chromosome and gene mapping. This sequence encodes a novel
				CC	human secreted and transmembrane PRO polypeptide
				XX	
				SQ	Sequence 2683 BP; 751 A; 604 C; 596 G; 732 T; 0 U; 0 Other;
					Query Match 93.4%; Score 2485; DB 7; Length 2683;
					Best Local Similarity 97.8%; Pred. No. 0;
					Matches 2576; Conservative 0; Mismatches 20; Indels 39; Gaps 4;
				QY	28 AACCTTGGTGCCTGCGTCCGTGAGAATTCAGCATGGAATGTCTCTACTATTTCTCTGGGAT 87
				Db	50 ACCTTGAGTGCCTGCGTCCGTGAGAATTCAGCATGGAATGTCTCTACTATTTCTCTGGGAT 109
				QY	88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAACGATTTTCATGATGTCTGG 147
				Db	110 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAACGATTTTCATGATGTCTGG 169
				QY	148 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGCTTCTG 207
				Db	170 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGCTGCTTCTG 229

QY 208 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 267
Db |||||
230 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 289
QY 268 AAAAATCTCTGGAAGGGAGGCCGTGTGCAGGCGGTCTCTGACCACTGACTCACCAGCCCTCG 327
Db |||||
290 AAAAATCTCTGGAAGGGAGGCCGTGTGCAGGCGGTCTCTGACCACTGACTCACCAGCCCTCG 349
QY 328 TGGGCTCAAAATATAACATTTGCGGTGAACCTGATATCCCTAGATGCCAAAGGAAGATG 387
Db |||||
350 TGGGCTCAAAATATAACATTTGCGGTGAACCTGATATCCCTAGATGCCAAAGGAAGATG 409
QY 388 CCAATGGCAACATAGTCTATGAGAAAGAACTGCAGAAATGAGGCTGGTTTTATCTGCTGATC 447
Db |||||
410 CCAATGGCAACATAGTCTATGAGAAAGAACTGCAGAAATGAGGCTGGTTTTATCTGCTGATC 469
QY 448 CATATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGACGGGAAAAATGGCACCGGCC 507
Db |||||
470 CGTATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGACGGGAAAAATGGCACCGGCC 529
QY 508 AAAGCCATCATAAACGTCTTCCCTGATGGGAAACCTTTTCTCACCAACCCCGGATGGAGAA 567
Db |||||
530 AAAGCCATCATAAACGTCTTCCCTGATGGGAAACCTTTTCTCACCAACCCCGGATGGAGAA 589
QY 568 GATGGAATTTCACTACGTCTTCCACACACACTTGGTCAGTATTTCCAGAAATTTGGACGAT 627
Db |||||
590 GATGGAATTTCACTACGTCTTCCACACACACTTGGTCAGTATTTCCAGAAATTTGGACGAT 649
QY 628 GTTCAGTGAGAGTTTCTGTGAACACAGCCAATGTGACACTTGGGCCTCAACTCATGGAAG 687
Db |||||
650 GTTCAGTGAGAGTTTCTGTGAACACAGCCAATGTGACACTTGGGCCTCAACTCATGGAAG 709
QY 688 TGACTGTCTACAGAAAGACATGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 747
Db |||||
710 TGACTGTCTACAGAAAGACATGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 769
QY 748 ACGTGGTAACAGATCAGATTCCTGTGTTTGTGACTATGTTCCAGAAACGATCGAAAT 807
Db |||||
770 ACGTGGTAACAGATCAGATTCCTGTGTTTGTGACTATGTTCCAGAAACGATCGAAAT 829
QY 808 CATCCGACGAAACCTTCC-CAAAGATCTCCCCATTAATGTTTGTGACTATGTTCCAGAAACGATCGAAAT 866
Db |||||
830 CATCCGACGAAACCTTCCCAAAGATCTCCCCATTAATGTTTGTGACTATGTTCCAGAAACGATCGAAAT 889
QY 867 CTAGCCACTTCCCTCAATTAATCTACCAATTAACATACTAAGTGGAGCTTCGGGGATAACTG 926
Db |||||
890 CTAGCCACTTCCCTCAATTAATCTACCAATTAACATACTAAGTGGAGCTTCGGGGATAACTG 949
QY 927 GCCTGTTGTTTCCACCAATCATACTGTGAATCACACGTATGTGCTCAATGGAACCTTCA 986
Db |||||
950 GCCTGTTGTTTCCACCAATCATACTGTGAATCACACGTATGTGCTCAATGGAACCTTCA 1009
QY 987 GCCTTAACCTCACTGTGAAAGCTGCAGCACGAGACCTTGTCCGCCACCGCCACCACCCAC 1046
Db |||||
1010 GCCTTAACCTCACTGTGAAAGCTGCAGCACGAGACCTTGTCCGCCACCGCCACCACCCAC 1069
QY 1047 CCAGACCTTCAAAACCCACCCCTTCTTT----- 1074
Db |||||
1070 CCAGACCTTCAAAACCCACCCCTTCTTTTAGCAACTACTTAAATCTTATGATTCAAACA 1129
QY 1075 ---AGGACCTGTGTGACAAACCCCTCGAGCTGAGTAGGATTCCTGATGAAAACTGCC 1130
Db |||||
1130 CCCAGGACCTTACTGTGACAAACCCCTCGAGCTGAGTAGGATTCCTGATGAAAACTGCC 1189
QY 1131 AGATTAAACAGATATGSCCACTTTCAAGCCACCATCACAAATGTAGAGGGGAATCTTAGAGG 1190
Db |||||
1190 AGATTAAACAGATATGSCCACTTTCAAGCCACCATCACAAATGTAGAGGGGAATCTTAGAGG 1249
QY 1191 TTAACATCATCCAGATGACAGACGTCTGTATGCGCGTGCCATGGCCTGAAAGCTCCCTAA 1250
Db |||||
1250 TTAACATCATCCAGATGACAGACGTCTGTATGCGCGTGCCATGGCCTGAAAGCTCCCTAA 1309

QY 1251 TAGACTTTGTCTGTGACCTGCCAAGGAGCATTTCCACGGAGGTCTGTACCATCATTTCTG 1310
Db |||||
1310 TAGACTTTGTCTGTGACCTGCCAAGGAGCATTTCCACGGAGGTCTGTACCATCATTTCTG 1369
QY 1311 ACCCACTCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT 1370
Db |||||
1370 ACCCACTCTGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT 1429
QY 1371 GTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCC 1430
Db |||||
1430 GTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCC 1489
QY 1431 TGGGGATGACACAAGCCCTGGCTCTCACGAGCACCCCTGATTTCTTCTTCTGACAGAGACC 1490
Db |||||
1490 TGGGGATGACACAAGCCCTGGCTCTCACGAGCACCCCTGATTTCTTCTTCTGACAGAGACC 1549
QY 1491 CAGCCTCGCCTTTAAGGATGGCAACACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATAT 1550
Db |||||
1550 CAGCCTCGCCTTTAAGGATGGCAACACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATAT 1609
QY 1551 TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAGAATAACAACCAATAGAAA 1610
Db |||||
1610 TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACACAGAATAACAACCAATAGAAA 1669
QY 1611 ATAGTCTGGGAATGTGTCAGAAAGCAAGGCTGAGTGCTTTTCTCAACCGTGCAAAAG 1670
Db |||||
1670 ATAGTCTGGGAATGTGTCAGAAAGCAAGGCTGAGTGCTTTTCTCAACCGTGCAAAAG 1729
QY 1671 CCGTGTCTTCCCGGGAACCAAGGATCCGCTACTCAAAAAACCAAGAAATTTAAAG 1730
Db |||||
1730 CCGTGTCTTCCCGGGAACCAAGGATCCGCTACTCAAAAAACCAAGAAATTTAAAG 1789
QY 1731 GAGTTTCTTAAATTTTGACCTTGTCTGAAAGCTCACTTTTCAAGTGCATTTGATGTGAGA 1790
Db |||||
1790 GAGTTTCTTAAATTTTGACCTTGTCTGAAAGCTCACTTTTCAAGTGCATTTGATGTGAGA 1849
QY 1791 TGTCTGGAGTGGCTATTAACCTTTTTTCTTCTTAAAGATTAATGTTAAATAGATATTGTGG 1850
Db |||||
1850 TGTCTGGAGTGGCTATTAACCTTTTTTCTTCTTAAAGATTAATGTTAAATAGATATTGTGG 1909
QY 1851 TTTGGGAAGTTGAATTTTATAGGTTAAATGTCAATTTTAGAGATGGGGAGAGGATTA 1910
Db |||||
1910 TTTGGGAAGTTGAATTTTATAGGTTAAATGTCAATTTTAGAGATGGGGAGAGGATTA 1969
QY 1911 TACTGCAGGCAGCTTCAGCCATGTTGTGAAACTGATAAAAGCAACTTAGCAAGGCTTCTT 1970
Db |||||
1970 TACTGCAGGCAGCTTCAGCCATGTTGTGAAACTGATAAAAGCAACTTAGCAAGGCTTCTT 2029
QY 1971 TTCATTATTTTATGTTTTCACTTATAAAGTCTTAGGTAACTAGTAGGATAGAAACACTG 2030
Db |||||
2030 TTCATTATTTTATGTTTTCACTTATAAAGTCTTAGGTAACTAGTAGGATAGAAACACTG 2089
QY 2031 TGTCCCGAGAGTAAGGAGAGAAGCTACTATTGATTAGAGCCTAACCCAGSTTAACTGCAA 2090
Db |||||
2090 TGTCCCGAGAGTAAGGAGAGAAGCTACTATTGATTAGAGCCTAACCCAGSTTAACTGCAA 2149
QY 2091 GAAGAGCGGGATACCTTTCAGCTTTCCATGTAACCTGTATGCATAAAGCCAATGTAGTCCA 2150
Db |||||
2150 GAAGAGCGGGATACCTTTCAGCTTTCCATGTAACCTGTATGCATAAAGCCAATGTAGTCCA 2209
QY 2151 GTTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTCC 2210
Db |||||
2210 GTTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTCC 2269
QY 2211 TGATGGAACAAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGTAGACTCAGAA 2270
Db |||||
2270 TGATGGAACAAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGTAGACTCAGAA 2329
QY 2271 AAAATACTACTCTCATAAATGGGTGGAGTATTTTGGTGACAACCTACTTTGCTTGGCTG 2330
Db |||||
2330 AAAATACTACTCTCATAAATGGGTGGAGTATTTTGGTGACAACCTACTTTGCTTGGCTG 2389
QY 2331 AGTGAAGGAATGATATTCATATATTTCATTTATTCCATGGACATTTAGTTAGTCTTTTAA 2390

QY 1047 CCAGACCTTCAAAACCCACCCCTTCTTT----- 1074
Db 1070 CCAGACCTTCAAAACCCACCCCTTCTTTAGCAACTACTCTFAAAATCTTATGATTCAAACA 1129
QY 1075 ---AGGACCTGCTGGTGACAAACCCCTCGAGCTGAGTAGGATTCCTGATGAAAACTGCC 1130
Db 1130 CCCAGGACCTACTGGTGACAACCCCTCGAGCTGAGTAGGATTCCTGATGAAAACTGCC 1189
QY 1131 AGATTAAACAGATATGGCCACTTTCAAGCCACCATCACAAATGTAGAGGGAATCTTAGAGG 1190
Db 1190 AGATTAAACAGATATGGCCACTTTCAAGCCACCATCACAAATGTAGAGGGAATCTTAGAGG 1249
QY 1191 TTAACATCATCCAGATGACAGACGTCCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAA 1250
Db 1250 TTAACATCATCCAGATGACAGACGTCCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAA 1309
QY 1251 TAGACTTTGTGTCGTGACCTGCCAAGGGAGCATTCOCACGGAGGTCTGTACCATCAATTCGTG 1310
Db 1310 TAGACTTTGTGTCGTGACCTGCCAAGGGAGCATTCOCACGGAGGTCTGTACCATCAATTCGTG 1369
QY 1311 ACCCCACCTGCGAGATCACCCAGAAACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT 1370
Db 1370 ACCCCACCTGCGAGATCACCCAGAAACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT 1429
QY 1371 GTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGACGTACTGTGTGAACCTCACCC 1430
Db 1430 GTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGACGTACTGTGTGAACCTCACCC 1489
QY 1431 TGGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCCCTGACAGAGACC 1490
Db 1490 TGGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCCCTGACAGAGACC 1549
QY 1491 CAGCCTCGCCTTTAAGGATGGCAACAGTGCCCTGATCTCCGTTGGCTGCTGGCCATAT 1550
Db 1550 CAGCCTCGCCTTTAAGGATGGCAACAGTGCCCTGATCTCCGTTGGCTGCTGGCCATAT 1609
QY 1551 TTGTCACTGTGATCTCCCTCTTGGTGTAACAAAAACAAGGAATACAAACCCCAATAGAAA 1610
Db 1610 TTGTCACTGTGATCTCCCTCTTGGTGTAACAAAAACAAGGAATACAAACCCCAATAGAAA 1669
QY 1611 ATAGTCCTGGGAATGTGGTCAGAAAGCAAGGCCTGAGTGTCTTTCTCAACCGTCGCAAAAG 1670
Db 1670 ATAGTCCTGGGAATGTGGTCAGAAAGCAAGGCCTGAGTGTCTTTCTCAACCGTCGCAAAAG 1729
QY 1671 CCGTGTCTTCCGGGAAACCCAGGAAAGGATCCGCTACTCAAAAACCAAGAATTTAAAG 1730
Db 1730 CCGTGTCTTCCGGGAAACCCAGGAAAGGATCCGCTACTCAAAAACCAAGAATTTAAAG 1789
QY 1731 GAGTTTCTTAAATTCGACCTTGTCTCTGAAGCTCACATTTTCAGTGCCCATGTGTGAGA 1790
Db 1790 GAGTTTCTTAAATTCGACCTTGTCTCTGAAGCTCACATTTTCAGTGCCCATGTGTGAGA 1849
QY 1791 TGTGCTGGAGTGGCTATTAAACCTTTTTTCTTAAAGATTATTGTTAAATAGATATTGTGG 1850
Db 1850 TGTGCTGGAGTGGCTATTAAACCTTTTTTCTTAAAGATTATTGTTAAATAGATATTGTGG 1909
QY 1851 TTTGGGGAAGTTGAATTTTTTATAGGTTAAATGTTCATTTTAGAGATGGGAGAGGGATTA 1910
Db 1910 TTTGGGGAAGTTGAATTTTTTATAGGTTAAATGTTCATTTTAGAGATGGGAGAGGGATTA 1969
QY 1911 TACTGCAGGACGTTTCAGCCATGTTGTGAAACTGATAAAAGCAACTTAGCAAGGCTTCTT 1970
Db 1970 TACTGCAGGACGTTTCAGCCATGTTGTGAAACTGATAAAAGCAACTTAGCAAGGCTTCTT 2029
QY 1971 TTCATTATTTTATGTTTCACTTATAAAGCTTAGGTAACTAGTAAGCCTAACCCAGGTTAACTG 2030
Db 2030 TTCATTATTTTATGTTTCACTTATAAAGCTTAGGTAACTAGTAAGCCTAACCCAGGTTAACTG 2089
QY 2031 TGTCCCGAGAGTAAGGAGAGAAAGCTACTATTGATTAGAGCCTAACCCAGGTTAACTGCAA 2090
Db 2090 TGTCCCGAGAGTAAGGAGAGAAAGCTACTATTGATTAGAGCCTAACCCAGGTTAACTGCAA 2149

QY 2091 GAAAGAGCGGGGATACTTTTCAGCTTTCCATGTAACCTGTAIGCATAAAGCCAATGTAGTCCA 2150
Db 2150 GAAAGAGCGGGGATACTTTTCAGCTTTCCATGTAACCTGTAIGCATAAAGCCAATGTAGTCCA 2209
QY 2151 GTTCTTAAGATCATGTTCCAAAGCTAACTGAATCCCACCTTCAATACACACTCATGAACCTCC 2210
Db 2210 GTTCTTAAGATCATGTTCCAAAGCTAACTGAATCCCACCTTCAATACACACTCATGAACCTCC 2269
QY 2211 TGATGGAACAATAACAGSCCCCAAGCCTGTGGTATGATGTGCACACTTGCTAGACTCAGAA 2270
Db 2270 TGATGGAACAATAACAGSCCCCAAGCCTGTGGTATGATGTGCACACTTGCTAGACTCAGAA 2329
QY 2271 AAAATACTACTCTCATAAATGGGTGGGAGTATTTTGGTGACAACCTACTTTGCTTGGCTG 2330
Db 2330 AAAATACTACTCTCATAAATGGGTGGGAGTATTTTGGTGACAACCTACTTTGCTTGGCTG 2389
QY 2331 AGTGAAGGAATGATATTCATATATATTCATTTATCCATGGACATTTAGTTAGTCTTTTTA 2390
Db 2390 AGTGAAGGAATGATATTCATATATATTCATTTATCCATGGACATTTAGTTAGTCTTTTTA 2449
QY 2391 TATACCAGGCATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTATAGTCG 2450
Db 2450 TATACCAGGCATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTACAGTCG 2509
QY 2451 CTGCACATATTTGAAATCAAAATATTAAGACTTTCCAAAAATTTGGTCCCTGGTTTTTCA 2510
Db 2510 CTGCACATATTTGAAATC-ATATATTAAGACTTTCCAAAGATGAGGTCCCTGGTTTTTCA 2568
QY 2511 TGGCAACTTGATCAGTAAGGATTTCCCTCTCTGTTGGAACTAAACCATTTACTATATGT 2570
Db 2569 TGGCAACTTGATCAGTAAGGATTTCCCTCTCTGTTGTAACCTAAACCATCTACTATATGT 2628
QY 2571 TAGACAAGACATTTTTTTTTTTTCTCTTCCCTTCCCTGAAAAA-AAAATGAGGGAAGAGACA 2624
Db 2629 TAGACATGACATTCCTTTTCTCTCTCTCTGAAAAAATAAAGTGTGGGAAGAGACA 2683

RESULT 13
ABT44529

ID ABT44529 standard; cDNA; 2683 BP.

XX ABT44529;

AC AC

XX 06-NOV-2003 (first entry)

DT Human PRO9925 cDNA.

XX

DE

XX

KW PRO; proliferation; gene; pericyte cell; TNF alpha; chondrocyte; blood;

KW tumour necrosis factor; proliferation; differentiation; gene therapy;

KW dermal fibroblast; ss.

XX

OS Homo sapiens.

XX

PN US2003027988-A1.

XX

PD 06-FEB-2003.

XX

PF 26-AUG-2002; 2002US-00227884.

XX

PR 01-JUN-2001; 2001WO-US017800.

PR 29-JUN-2001; 2001WO-US021066.

PR 09-APR-2002; 2002US-00119480.

XX

PA (GETH) GENENTECH INC.

XX

PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;

PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX

DR WPI; 2003-503301/47.

DR P-PSDB; ABJ72376.

XX

PT New PRO protein encoding nucleic acid, useful for preparing PRO

polypeptides and anti-PRO antibodies for detecting the presence of a

Db 1910 TTTGGGAAGTTGAATTTTATAGGTTAAATGTCATTTTAGAGATGGGAGAGGGATTA 1969
QY 1911 TACTGCAGGCAGCTTCAGCCATGTTGTGAACTGATAAAAGCACTTAGCAAGGCTTCTT 1970
Db 1970 TACTGCAGGCAGCTTCAGCCATGTTGTGAACTGATAAAAGCACTTAGCAAGGCTTCTT 2029
QY 1971 TTCATTATTTTATGTTTCACTTATAAGTCTTAGGTAACTAGTAGATAGAAAACACTG 2030
Db 2030 TTCATTATTTTATGTTTCACTTATAAGTCTTAGGTAACTAGTAGATAGAAAACACTG 2089
QY 2031 TGTCCCGAGAGTAAGGAGAGAGAGCTACTATTGATTAGAGCCTAACCCAGGTTAACTGCAA 2090
Db 2090 TGTCCCGAGAGTAAGGAGAGAGAGCTACTATTGATTAGAGCCTAACCCAGGTTAACTGCAA 2149
QY 2091 GAAGAGCGGGATACCTTTCAGCTTTCATGTAACTGTATGCATAAAGCCAAATGTAGTCCA 2150
Db 2150 GAAGAGCGGGATACCTTTCAGCTTTCATGTAACTGTATGCATAAAGCCAAATGTAGTCCA 2209
QY 2151 GTTCTAAGATCATGTTTCCAAGCTAACTGAATCCCACTTCAATACACACTCATGAATCC 2210
Db 2210 GTTCTAAGATCATGTTTCCAAGCTAACTGAATCCCACTTCAATACACACTCATGAATCC 2269
QY 2211 TGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGCTAGACTCAGAA 2270
Db 2270 TGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGCTAGACTCAGAA 2329
QY 2271 AAAATACTACTCTCATAAATGGGTGGGAGTATTTTGGTGACAACCTACTTTTGCTGGCTG 2330
Db 2330 AAAATACTACTCTCATAAATGGGTGGGAGTATTTTGGTGACAACCTACTTTTGCTGGCTG 2389
QY 2331 AGTGAAGGAATGATATTCATATATTCATTTATCCATGGACATTTAGTAGTCTTTTA 2390
Db 2390 AGTGAAGGAATGATATTCATATATTCATTTATCCATGGACATTTAGTAGTCTTTTA 2449
QY 2391 TATACCAGGCATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTATAGTCG 2450
Db 2450 TATACCAGGCATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTACAGTCG 2509
QY 2451 CTGCACATATTTGAAATCAAAATATTAAGACTTTCAAAAATTTGGTCCCTGGTTTTC 2510
Db 2510 CTGCACATATTTGAAATC-ATATATTAAGACTTTCAAAAATGAGGTCCCTGGTTTTC 2568
QY 2511 TGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTGGAACTAAACCATTACTATATGT 2570
Db 2569 TGGCAACTTGATCAGTAAGGATTTCACTCTGTTTGTAACTAAACCATTACTATATGT 2628
QY 2571 TAGACAAGACATTTTCTTCTTCTTCTGTAATTTCCCTGTAATAA-AAAATGAGGGAAGAGACA 2624
Db 2629 TAGACATGACATTTCTTCTTCTTCTGTAATAAATAAGTGTGGGAAGAGACA 2683

RESULT 14
ACD82196
ID ACD82196 standard; cDNA; 2683 BP.
XX
AC ACD82196;
XX
DT 19-SEP-2003 (first entry)
XX
DE Human secreted/transmembrane polypeptide PRO 9925 cDNA.
XX
KW Human; ss; chondrocyte stimulation; TNF-alpha stimulation; gene therapy;
KW human dermal fibroblast stimulation; tumour; tissue typing; gene;
KW affinity purification.
XX
OS Homo sapiens.
XX
PN US2003044934-A1.
XX
PD 06-MAR-2003.
XX
PF 28-AUG-2002; 2002US-00230338.

XX
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
PA (GETH) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
DR WPI; 2003-492274/46.
DR P-PSDB; ABO34271.
XX
PT New transmembrane polypeptides and nucleic acids encoding the
PT polypeptides, useful in gene therapy, in chromosome identification, as
PT chromosome markers, or in generating probes.
XX
PS Claim 2; Fig 41; 315pp; English.
XX
CC The invention relates to an isolated nucleic acid encoding a PRO
CC polypeptide. Nucleic acids that encode PRO can be used to generate either
CC transgenic animals or knock-out animals useful in developing and
CC screening of therapeutically useful reagents. The nucleic acids may also
CC be used in gene therapy for replacing defective gene, in chromosome
CC identification, as chromosome markers, or in generating probes to isolate
CC full length PRO cDNA. The PRO polypeptides are useful for chondrocyte
CC stimulation, TNF-alpha stimulation, human dermal fibroblasts stimulation
CC and for detecting the presence of tumour in an mammal. The PRO
CC and the isolated nucleic acids may be used for protein electrophoresis
CC those markers. The PRO polypeptides and nucleic acids may also be used in
CC tissue typing. Anti-PRO antibodies are useful in diagnostic assays for
CC PRO and in affinity purification of PRO from recombinant cell culture or
CC natural sources. The present sequence represents cDNA encoding a human
CC secreted/transmembrane PRO polypeptide
XX
SQ Sequence 2683 BP; 751 A; 604 C; 596 G; 732 T; 0 U; 0 Other;

Query Match 93.4%; Score 2485; DB 8; Length 2683;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 2576; Conservative 0; Mismatches 20; Indels 39; Gaps 4;

QY 28 AACCTTGGTGCTGCTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCTGGGAT 87
Db 50 ACCTTGAGTGCTGCTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCTGGGAT 109
QY 88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCGCGCCAAACGATTTTCATGATGTCTGG 147
Db 110 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCGCGCCAAACGATTTTCATGATGTCTGG 169
QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGAGACAAATCAATTAATGGCTGGTCTTCTG 207
Db 170 GCAATGAAAGACCTTCTGCTTACATGAGGAGACAAATCAATTAATGGCTGGTCTTCTG 229
QY 208 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 267
Db 230 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGGA 289
QY 268 AAAACTCCTGGAAGGGAGCGCTGTGCAGCGCGTCTGACCAGTGACTCACCAGCCCTCG 327
Db 290 AAAACTCCTGGAAGGGAGCGCTGTGCAGCGCGTCTGACCAGTGACTCACCAGCCCTCG 349
QY 328 TGGGCTCAAATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG 387
Db 350 TGGGCTCAAATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG 409
QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGTGGTTATCTGTGATC 447
Db 410 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGTGGTTATCTGTGATC 469
QY 448 CATATGTTTACAACTGGACAGCATGGTCAGAGGACAGTACGCGGAAAAATGGCACCGGCC 507
Db 470 CGTATGTTTACAACTGGACAGCATGGTCAGAGGACAGTACGCGGAAAAATGGCACCGGCC 529

QY	508	AAAGCCATCATAAACGCTCTCCCTGATGGGAAACCTTTTCTCTCACCACCCCGGATGGAGAA	567
Db	530	AAAGCCATCATAAACGCTCTCCCTGATGGGAAACCTTTTCTCTCACCACCCCGGATGGAGAA	589
QY	568	GATGGAATTTCAATCTACGCTCTCCACACACTTTGGTCAGTATTTCCAGAAATTTGGACGAT	627
Db	590	GATGGAATTTCAATCTACGCTCTCCACACACTTTGGTCAGTATTTCCAGAAATTTGGACGAT	649
QY	628	GTTTCAGTGAGAGTTTCTGTGAACACAGCCAATGTGACACTTGGGCTCAACTCATGGAAG	687
Db	650	GTTTCAGTGAGAGTTTCTGTGAACACAGCCAATGTGACACTTGGGCTCAACTCATGGAAG	709
QY	688	TGACTGTCTACAGAAAGACATGACGCGGCATATGTTCCCATCGCACAAAGTGAAAGATGTGT	747
Db	710	TGACTGTCTACAGAAAGACATGACGCGGCATATGTTCCCATCGCACAAAGTGAAAGATGTGT	769
QY	748	ACGTGGTAAACAGATCAGATTCCTGTGTTTGTGACTATGTTCCAGAAACGATCGAAATTT	807
Db	770	ACGTGGTAAACAGATCAGATTCCTGTGTTTGTGACTATGTTCCAGAAACGATCGAAATTT	829
QY	808	CATCCGACGAAACCTTCC-CAAAAGATCTCCCCATTAATGTTGATGTCCTGATTCATGATC	866
Db	830	CATCCGACGAAACCTTCCCAAAGATCTCCCCATTAATGTTGATGTCCTGATTCATGATC	889
QY	867	CTAGCCACTTCCCTCAATTAATCTACCAATTAACACTACAAGTGGAGTTTCGGGGATAATACTG	926
Db	890	CTAGCCACTTCCCTCAATTAATCTACCAATTAACACTACAAGTGGAGTTTCGGGGATAATACTG	949
QY	927	GCCTGTTGTTTCCACCAATCATCTGTGAATCACACGATATGTCCTCAATGGAACCTTCA	986
Db	950	GCCTGTTGTTTCCACCAATCATCTGTGAATCACACGATATGTCCTCAATGGAACCTTCA	1009
QY	987	GCCTTAACCTCACTGTGAAAGCTGCAGCACGAGGACCTTGTCCGCCACCCACCCAC	1046
Db	1010	GCCTTAACCTCACTGTGAAAGCTGCAGCACGAGGACCTTGTCCGCCACCCACCCAC	1069
QY	1047	CCAGACCTTCAAAACCCACCCCTTCTTT	1074
Db	1070	CCAGACCTTCAAAACCCACCCCTTCTTTAGCAACTACTTAAATCTTATGATTCAAACA	1129
QY	1075	---AGGACCTGCTGGTGACAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACCTGCC	1130
Db	1130	CCCCAGGACCTACTGCTGACAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACCTGCC	1189
QY	1131	AGATTAAACAGATATGGCCACTTTTCAAGCCACCATCACAATTTGAGAGGGAATCTTAGAGG	1190
Db	1190	AGATTAAACAGATATGGCCACTTTTCAAGCCACCATCACAATTTGAGAGGGAATCTTAGAGG	1249
QY	1191	TTAACATCATCCAGATGACAGACGTCCTGATGCGGTCCTATGGCCTGAAAGCTCCCTAA	1250
Db	1250	TTAACATCATCCAGATGACAGACGTCCTGATGCGGTCCTATGGCCTGAAAGCTCCCTAA	1309
QY	1251	TAGACTTTGTCGTGACTGCCAAGGGAGCAATCCACGGAGGTCTGTACCATCATTTCTG	1310
Db	1310	TAGACTTTGTCGTGACTGCCAAGGGAGCAATCCACGGAGGTCTGTACCATCATTTCTG	1369
QY	1311	ACCCACCTGCGAGATCACCCAGAAACACAGTCGTGAGCCCTGTGGATGTGATGATGT	1370
Db	1370	ACCCACCTGCGAGATCACCCAGAAACACAGTCGTGAGCCCTGTGGATGTGATGATGATGT	1429
QY	1371	GTCTGCTGACTGTGAGACGAACCTTCAATGGGTCGTGGACGTCATGTGTGAACCTCACCC	1430
Db	1430	GTCTGCTGACTGTGAGACGAACCTTCAATGGGTCGTGGACGTCATGTGTGAACCTCACCC	1489
QY	1431	TGGGGGATGACACAAGCCTGGCTCTCACGAGCACCTGATTTCTGTTCTGACAGAGACC	1490
Db	1490	TGGGGGATGACACAAGCCTGGCTCTCACGAGCACCTGATTTCTGTTCTGACAGAGACC	1549
QY	1491	CAGCCTCGCCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATAT	1550
Db	1550	CAGCCTCGCCTTTAAGGATGGCAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATAT	1609

Db	1130	CCCCAGGACCTACTGGTGACAAACCCCTGGAGCTGAGTAGGATTCTCTGATGAAAACTGCC	1189	2210	GTTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCACCTTCAATACACACTCATGAACCTCC	2265
QY	1131	AGATTAAACAGATATGSCCACTTTCAAGCCACCATCACAAATTGTAGAGGGAATCTTAGAGG	1190	2211	TGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGTAGACTCAGAA	2270
Db	1190	AGATTAAACAGATATGSCCACTTTCAAGCCACCATCACAAATTGTAGAGGGAATCTTAGAGG	1249	2270	TGATGGAACAATAACAGGCCCAAGCCTGTGGTATGATGTGCACACTTGTAGACTCAGAA	2329
QY	1191	TTAACATCATCCAGATGACAGACGTCCTGTATGCCGCTGCCATGGCCTGAAAGCTCCCTAA	1250	2271	AAAATACTACTCTCATAAATGGTGGGAGTATTTTGGTGACAACCTACTTTGCTTGGCTG	2330
Db	1250	TTAACATCATCCAGATGACAGACGTCCTGTATGCCGCTGCCATGGCCTGAAAGCTCCCTAA	1309	2330	AAAATACTACTCTCATAAATGGTGGGAGTATTTTGGTGACAACCTACTTTGCTTGGCTG	2389
QY	1251	TAGACTTTTGTGTCGTAACCTGCCAAGGGAGCATTTCCACGGAGGTCCTGTACCATCATTTCTG	1310	2331	AGTGAAGGAATGATATTTCATATATTTCATTTATTTCCATGGACATTTAGTTAGTCTTTT	2390
Db	1310	TAGACTTTTGTGTCGTAACCTGCCAAGGGAGCATTTCCACGGAGGTCCTGTACCATCATTTCTG	1369	2390	AGTGAAGGAATGATATTTCATATATTTCATTTATTTCCATGGACATTTAGTTAGTCTTTT	2449
QY	1311	ACCCACCTGCGAGATCACCCAGAAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT	1370	2391	TATACAGGCATGATGCTGAGTGACACTCTTGTGTATATTTTCCAAATTTTGTATAGTCG	2450
Db	1370	ACCCACCTGCGAGATCACCCAGAAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGT	1429	2450	TATACAGGCATGATGCTGAGTGACACTCTTGTGTATATTTTCCAAATTTTGTACAGTCG	2509
QY	1371	GTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCC	1430	2451	CTGCACATATTTGAAATCAAAATATTAAGACTTTCAAAAAATTTGTCCTGGTTTTTCA	2510
Db	1430	GTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCC	1489	2510	CTGCACATATTTGAAATC-ATATATTAAGACTTTCCAAAGATGAGTCCCTGGTTTTTCA	2568
QY	1431	TGGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCCCTGACAGAGCC	1490	2511	TGGCACTTGATCAGTAAGGATTTCCCTCTGTGTTTGGAACTAAACCATTACTATATGT	2570
Db	1490	TGGGGGATGACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCCCTGACAGAGCC	1549	2569	TGGCACTTGATCAGTAAGGATTTTCACCTCTGTTTGTAACTAAACCATTACTATATGT	2628
QY	1491	CAGCCTCGCCTTTAAGGATGGCAACAGTGCCCTGATCTCCGTTGGCTGCTTGGCCATAT	1550	2571	TAGACAAGACATTTTTTTTTTTTTTTCCTTCCTGAAAAA-AAAAATGAGGGAAGAGACA	2624
				2629	TAGACATGACATTTCTTTTCTCTCTCTCTGAAAAATAAAGCTGTGGGAAGAGACA	2683

Search completed: September 10, 2004, 02:34:51
Job time : 1498 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 10, 2004, 15:48:34 ; Search time 569.5 Seconds
(without alignments)
2996.857 Million cell updates/sec

Title: US-10-039-272-1
Perfect score: 4850
Sequence: 1 cggcacgagggcccgagga.....aaaaaaaaaaaaaaaa 2661

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 2670352

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10039272/runat_07092004_144225_21764/app query.fasta_1.2823
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosu62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10039272@cgn_1_1_262@runat_07092004_144225_21764
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

#			
Result No.	Score	Query Match Length	Description

1	2999	61.8	560	9	US-09-735-705-225	Sequence 225, App
2	2999	61.8	560	9	US-09-850-716A-225	Sequence 225, App
3	2999	61.8	560	9	US-09-897-778-225	Sequence 225, App
4	2999	61.8	560	9	US-09-943-075A-6	Sequence 6, Appli
5	2999	61.8	560	12	US-10-007-700-225	Sequence 225, App
6	2999	61.8	560	12	US-10-463-106-2	Sequence 2, Appli
7	2999	61.8	560	13	US-10-039-272-2	Sequence 2, Appli
8	2999	61.8	560	14	US-10-117-982-225	Sequence 225, App
9	2999	61.8	560	15	US-10-295-027-1258	Sequence 1258, Ap
10	2999	61.8	560	15	US-10-313-986-225	Sequence 225, App
11	2999	61.8	560	16	US-10-309-290-152	Sequence 152, App
12	2999	61.8	560	16	US-10-408-765A-466	Sequence 466, App
13	2979	61.4	572	12	US-10-219-535-42	Sequence 42, Appl
14	2979	61.4	572	12	US-10-232-230-42	Sequence 42, Appl
15	2979	61.4	572	12	US-10-232-224-42	Sequence 42, Appl
16	2979	61.4	572	14	US-10-227-884-42	Sequence 42, Appl
17	2979	61.4	572	14	US-10-230-163-42	Sequence 42, Appl
18	2979	61.4	572	14	US-10-230-338-42	Sequence 42, Appl
19	2979	61.4	572	14	US-10-218-631-42	Sequence 42, Appl
20	2979	61.4	572	14	US-10-230-414-42	Sequence 42, Appl
21	2979	61.4	572	14	US-10-216-159A-42	Sequence 42, Appl
22	2979	61.4	572	14	US-10-218-849-42	Sequence 42, Appl
23	2979	61.4	572	14	US-10-227-873-42	Sequence 42, Appl
24	2979	61.4	572	14	US-10-227-883-42	Sequence 42, Appl
25	2979	61.4	572	14	US-10-219-076-42	Sequence 42, Appl
26	2979	61.4	572	14	US-10-230-434-42	Sequence 42, Appl
27	2979	61.4	572	14	US-10-219-003-42	Sequence 42, Appl
28	2979	61.4	572	14	US-10-219-075-42	Sequence 42, Appl
29	2979	61.4	572	14	US-10-219-464-42	Sequence 42, Appl
30	2979	61.4	572	14	US-10-219-466-42	Sequence 42, Appl
31	2979	61.4	572	14	US-10-219-479-42	Sequence 42, Appl
32	2979	61.4	572	14	US-10-219-481-42	Sequence 42, Appl
33	2979	61.4	572	14	US-10-230-260-42	Sequence 42, Appl
34	2979	61.4	572	14	US-10-232-231-42	Sequence 42, Appl
35	2979	61.4	572	14	US-10-232-233-42	Sequence 42, Appl
36	2979	61.4	572	14	US-10-216-165-42	Sequence 42, Appl
37	2979	61.4	572	14	US-10-218-956-42	Sequence 42, Appl
38	2979	61.4	572	14	US-10-219-468-42	Sequence 42, Appl
39	2979	61.4	572	14	US-10-219-478-42	Sequence 42, Appl
40	2979	61.4	572	14	US-10-219-536-42	Sequence 42, Appl
41	2979	61.4	572	14	US-10-233-205-42	Sequence 42, Appl
42	2979	61.4	572	14	US-10-219-072-42	Sequence 42, Appl
43	2979	61.4	572	14	US-10-219-470-42	Sequence 42, Appl
44	2979	61.4	572	14	US-10-219-474-42	Sequence 42, Appl
45	2979	61.4	572	14	US-10-219-524-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1
US-09-735-705-225
; Sequence 225, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735,705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-735-705-225

Alignment Scores:
Pred. No.: 7,23e-270 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservatives: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 9 Gaps: 0

US-10-039-272-1 (1-2661) x US-09-735-705-225 (1-560)

QY 60 ATGGAATGCTCTACTATTTCCTGGGATTTCCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAAATGGCTGCTTCTGATGAAATGACTGGAATGAAAACTCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTyrSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAATACTCTGTAAGGAGGCCGCTGTCAGGG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80
QY 300 GTCCTGACGAGTCACTCACCAGCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGGAATGGCACCGGCCAAAGCCATCATACGCTTCTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProAspGlyLys 160
QY 540 CTTTTCCTCACCACCCGGATGGAGAAGATGGAATTCATCTACGTCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAAGTATTTCCAGAAATTGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAACAGATCAGATTCTCTGTGTTTGG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTGATGCTCTGATTCATGATTCATGACCTAGCCACTTCTCCTCAATTATTCTACCAATTAA 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTCGGGGATAATACTGGCCTGTTTGTGTTTCCCAACATCATCTGTGAAT 958

Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCGGCCACCGCCACCCACCCAGACCTTCAAAACCCACCCCTCTTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProProProProProProSerLeuGly 340
QY 1079 CCTGCTGGTGACAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACTGCCAGATTAA 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTCAAGCCACCATCACAAATTGTAGAGGAATCTTAGAGGTTAACATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGTCTGATGCCGTGCCATGGCCTGAAAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACCTGCCAAGGGAGCATTCACCGAGGCTGTGTACCATCATTTCTGACCCACC 1318
Db 401 valValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TCGGAGATCACCAGAACACAGTCTGTGAGCCCTGTGGATGTGGATGAGATGTGTCTGTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACCTTCAATGGTCTGGGACGTACTGTGTGAACCTCACCTCGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACACAAGCCTGGCTCTCACGAGCACCCCTGATTCTGTTCCTGTGACAGAGCCAGCCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGGATGGCAAAACAGTCCCTGATCTCCGTGGCTGGCTGGCCATATTGTCACT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTTGTGTACAAAAACACAGGAATACAAACCAATAGAAAAATAGTCCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGGAAATGTGTCAGAAAGCCCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTTTC 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGGAAACCAGGAAAGGATCCGCTACTCAAAAAACCAAGAAATTTAAAGGAGTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 2

US-09-850-716A-225
; Sequence 225, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT

; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-225

Alignment Scores:
Pred. No.: 7.23e-270 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 9 Gaps: 0

US-10-039-272-1 (1-2661) x US-09-897-778-225 (1-560)
QY 60 ATGGAATGCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTCATGATGTGCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAAATGGCTGGTCTTCTGATGAAAATGACTGGAATGAAAAAATCTACCCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAAAAATCCTCTGGAAGGAGGCCGTGTGCAGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80
QY 300 GTCCTGACCACTCACTCACCAGCCCTCGTGGCTCAAATATAACATTTCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAATGGACAGCATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGGAATAATGGCACCGGCCAAAGCCATCATACCTCTTCCCTGATGGGAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACCACCCGGATGGAGAAGATGGAATTTTCATCTACGTCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAAGTATTTCCAGAAATGGGACGATGTTCAAGTGAAGTCTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTGAACAGATCAGATTCTGTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAGAACGATCGAAATTCATCCGACGGAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTGATGTCCTGATTCATGATCCTAGCCACTTCTCTCAATTATTCTACCATTAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTGTTTCCACCAATCATACTGTGAAT 958

Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTGCTCAATGGAACCTTTCAGCCCTTAACCTCACTGTGAAAGCTGCAGCACA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCGGCCACCGCCACACCACCCAGACCTTCAAAACCCACCCCTTCTTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProProProProProProSerLeuGly 340
QY 1079 CCTGCTGGTGACAACCCCTCGAGCTGAGTAGGATTCTCTGATGAAAACTGCCAGATTAAAC 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTCAAGCCACCATCACAAATTGTAGAGGGAATCTTAGAGGTTAACATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGTCTCTGATGCGGTGCCATGGCCTGAAAGTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACCTGCCAAGGAGCATTCACACGGAGGTCTGTACCATCATTTCTGACCCACC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TCGGAGATCACCCAGAACACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCTTGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCCTGACAGACCCAGCCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGGATGCAAAACAGTCCCTGATCTCCGTGGCTGGCTGCTTGGCCATATTGTCACT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTTGGTGTAACAAAAACACAGGAATACACCCCAATAGAAAAATAGTCCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGGAATGTGTCAGAACAAAGGCTGAGTCTTTCTCAACCCGTGCAAAAGCCGTGTTC 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGGAACCCAGGAAAAGGATCCGCTACTCAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 4
US-09-943-075A-6
; Sequence 6, Application US/09943075A
; Patent No. US20020151486A1
; GENERAL INFORMATION:
; APPLICANT: Popoff, Steven N.
; APPLICANT: Safado, Faye F.
; APPLICANT: Owen, Thomas A.
; APPLICANT: Smock, Steven L.
; TITLE OF INVENTION: Osteoactivin Protein and Nucleic Acids Encoding the Same,
; TITLE OF INVENTION: Compositions and Methods of Stimulating Bone Differentiation
; FILE REFERENCE: 71369.262
; CURRENT APPLICATION NUMBER: US/09/943,075A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: US 60/229,006
; PRIOR FILING DATE: 2000-08-30
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSEQ for Windows Version 4.0

; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Peckman, David W.
; APPLICANT: Cai, Feng
; APPLICANT: Foy, Teresa M.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C17
; CURRENT APPLICATION NUMBER: US/10/007,700
; CURRENT FILING DATE: 2001-11-30
; NUMBER OF SEQ ID NOS: 469
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-700-225

Alignment Scores:
Pred. No.: 7.23e-270 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 12 Gaps: 0

US-10-039-272-1 (1-2661) x US-10-007-700-225 (1-560)

QY 60 ATGGAATGCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAATGGCTGCTTCTGATGAAATGACTGGAATGAAAACTCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAAAACTCCTGGAAGGAGGCCGCTGTCAGGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80
QY 300 GTCCTGACCAGTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTTCGGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGAAATGGCACCGGCCAAAGCCCATCAACGCTCTTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACCACCCCGGATGGAGAGATGGAATTTTCATCTACGTCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAAGTATTTCCAGAAATTTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAGTGAAGATGTGTACGTGGTGAACAGATCAGATTCTCTGTGTTTGTG 779

Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTGATGCTCTGATTCATGATCCTAGCCACTTCTCCTCAATATTCTACCATTAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATCTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCGCCACCACGCCACCCAGACCTTCAAAACCCACCCCTTCTTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProProProProProProSerLeuGly 340
QY 1079 CCTGCTGTGACAAACCCCTGGAGCTGAGTAGGATTCTCTGATGAAAACTGCCAGATTAAAC 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGCCACTTTCAAGCCACCATCACAAATGTAGAGGGAATCTTAGAGGTTAACATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACAGTCCTGATGCCGCTGAGGCTGACCATCATCTTCTGACCCACCC 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACCTGCCAAGGAGCATTCGCCACGAGGTCTGTACCATCATCTTCTGACCCACCC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TGCAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGGACCTACTGTGTGAACCTCACCCCTGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACACAAGCCTGGCTCTCACGAGCACCTGATTCTGTCTCTGACAGAGACCCAGCCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGGATGGCAACAGTGCCTGATCTCCGTGGCTGCTTGGCCATATTGTCACT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTTGGTGTACAAAAACACAAAGGAATACAAACCAATAGAAAAATAGTCCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GCGAATGTGGTCAGAACCAAGCCCTGAGTGTCTTCTCAACCGTGCAGAAAGCCGTGTTTC 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGGAACACGAGAAAGGATCCGCTACTCAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 6

US-10-463-106-2
; Sequence 2, Application US/10463106
; Publication No. US20030202938A1
; GENERAL INFORMATION:
; APPLICANT: RAMESHWAR, Pranela
; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-1 GENE


```
; FILE REFERENCE: 267/033 University of Medicine & Dentistry of New Jersey
; CURRENT APPLICATION NUMBER: US/10/463,106
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US 10/039,272
; PRIOR FILING DATE: 2001-10-20
; PRIOR APPLICATION NUMBER: US 60/241,881
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-463-106-2

Alignment Scores:
Pred. No.:      7.23e-270      Length:      560
Score:          2999.00      Matches:      559
Percent Similarity: 99.82%      Conservative: 0
Best Local Similarity: 99.82%      Mismatches: 1
Query Match:      61.84%      Indels:      1
DB:              12          Gaps:      0

US-10-039-272-1 (1-2661) x US-10-463-106-2 (1-560)

QY 60 ATGGAATGTCCTACTATTTCCTGGGATTTCTGCTCCCTGGTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20

QY 120 GCCGCCAAACGATTTTCATGATGTCTGCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40

QY 180 CACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60

QY 240 GTGTGGAAGCGGGAGACATGAGGTGGAAAACTCCTGGAAGGGAGGCCGTGTGCAGGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80

QY 300 GTCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAAATATAACATTTGCGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100

QY 360 ATATTCCTAGATGCCAAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120

QY 420 AGAAATGAGGCTGGTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140

QY 480 GACAGTGACGGGAAAAATGGCACC GGCCCAAGCCCATCAACGTCCTTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160

QY 540 CCTTTTCCTCACCACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCCTCCACACACTT 599
Db 161 PropheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180

QY 600 GGTCAAGTATTCAGAAAATTTGGACGATGTTCACTGAGAGTTTCTGTGTAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200

QY 660 GTGACACTTGGGCTCAACTCATGGAAGTCACTGTCTACAGAAGACATGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220

QY 720 GTTCCCATCGCACAAAGTGAAAGATGTGTACGTGGTAAACAGATCAGATCCTGTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
```

```
QY 780 ACTATGTTCCAGAAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260

QY 839 ATTATGTTTGATGTCTGATTTCATGATCCTTAGCCACTTCCTCAATTATTCTACCATTAAC 898
Db IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280

QY 899 TACAAGTGGAGCTTCGGGGATAATACTAGGCCTGTTTGTTCACCAATCATACTGTGAAT 958
Db TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300

QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACCA 1018
Db HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320

QY 1019 GGACCTTGTCCGCCACCGCCACCAACCACCCAGACCTTCAAAAACCCACCCCTTCTTTAGGA 1078
Db GlyProCysProProProProProProProArgProSerLysProThrProSerLeuGly 340

QY 1079 CCTGCTGGTGACAACCCCTGGAGCTGAGTAGGATTCTCTGATGAAAACTGCCAGATTAAAC 1138
Db ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360

QY 1139 AGATATGGCCACTTTCAGGCACCAACCATCACAAATGTAGAGGGAATCTTAGAGGTTACATC 1198
Db ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380

QY 1199 ATCCAGATGACAGACGTCCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTT 1258
Db IleGlnMetThrAspValLeuMetProValProTyrProGluSerSerLeuIleAspPhe 400

QY 1259 GTCGTGACCTGCCAAGGAGCATTTCCCACGGAGGTCTGTACCATCATTTCTGACCCACC 1318
Db ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420

QY 1319 TGCAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTGCTGTG 1378
Db CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440

QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGGACGTA CTGTGAACCTCACCCCTGGGGAT 1438
Db ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460

QY 1439 GACACAAGCCTGCTCTCACGAGCACCCCTGATTTCTGTTCTTGACAGAGACCCAGCCTCG 1498
Db AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480

QY 1499 CCTTTAAGGATGGCAAAACAGTGGCCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACT 1558
Db ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500

QY 1559 GTGATCTCCCTCTTGGTGTA CA AAAACACAAGGAATACACCCCAATAGAAAATAGTCCT 1618
Db ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520

QY 1619 GGAATGTGTCAGAAAGCAAGGCCCTGAGTGTCTTTCTCAACCCGTGCAAAAAGCCGTGTT 1678
Db GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540

QY 1679 TTCCCGGGAACCCAGGAAAAGGATCCGCTACTCAAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560
```

```
RESULT 7
US-10-039-272-2
; Sequence 2, Application US/10039272
; Publication No. US20020168653A1
; GENERAL INFORMATION:
; APPLICANT: RAMESHWAR, Pranela
; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-TYPE
; FILE REFERENCE: 267/033 University of Medicine & Dentistry of New Jersey
; CURRENT APPLICATION NUMBER: US/10/039,272
```

```

; CURRENT FILING DATE: 2001-10-20
; PRIOR APPLICATION NUMBER: US 60/241,881
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 560
;

```

Alignment Scores:					
Pred. No.:	7.23e-270	Length:	560		
Score:	2999.00	Matches:	559		
Percent Similarity:	99.82%	Conservative:	0		
Best Local Similarity:	99.82%	Mismatches:	1		
Query Match:	61.84%	Indels:	1		
DB:	13	Gaps:	0		

US-10-039-272-1 (1-2661) x US-10-039-272-2 (1-560)

QY	60	ATGGAATGTCTCTACTATTTCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT	119
Db	1	MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp	20
QY	120	GCCGCCAAACGATTTTCATGATGTGCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAG	179
Db	21	AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu	40
QY	180	CACAATCAATTAAATGGCTGGTCTTCTGATGAAAAATGACTGGAAATGAAAAACTCTACCCA	239
Db	41	HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60
QY	240	GTGTGGAAGCGGGAGACATGAGGTGAAAAAACTCCTGGAAGGAGGCCGTGTGCAGGCG	299
Db	61	ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla	80
QY	300	GTCCTGACCAGTGACTCACCGCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTG	359
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100
QY	360	ATATTCCCTAGATGCCAAAAGGAGATGCCAATGGCAACATAGTCTATGAGAAAGAACTGC	419
Db	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120
QY	420	AGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG	479
Db	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140
QY	480	GCAGTGCAGGGGAAAAATGGCACCGGCCAAAGCCATCATAACGCTCTTCCCTGATGGAAA	539
Db	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys	160
QY	540	CCTTTTCCTCACCAACCCCGGATGGAGAGATGGAATTTTCATCTACGCTTCCACACACTT	599
Db	161	ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu	180
QY	600	GGTCAGTATTTCCAGAAAATTTGGACGATGTTTCAGTGAGAGTTTCTGTGAAACACAGCCAAT	659
Db	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200
QY	660	GTGACACTTGGGCCTCAACTCATGGAAGTGAAGTGTCTACAGAAGACATGGACGGGCATAT	719
Db	201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220
QY	720	GTTCCCATCGCACAAGTGAAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTTGTG	779
Db	221	ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal	240
QY	780	ACTATGTTCCAGAGAACGATCGAAATTTCATCCGACGAAACCTTCCC-AAAGATCTCCCC	838
Db	241	ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro	260

QY	839	ATTATGTTTGATGTCTCCTGATTCAATCATGATCCCTAGCCACTTCCTCCAATTATTTCTACCATTAAC	898
Db	261	IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280
QY	899	TACAAGTGGAGCTTCGGGGATAATACTGGCCCTGTTTGTTCACCAATCATACTGTGAAT	958
Db	281	TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300
QY	959	CACACGTATGTCTCAATGGAAACCTTCAGCCTTAACCTCACCTGTGAAAAGCTGCAGCACCA	1018
Db	301	HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro	320
QY	1019	GGACCTTGTCGCCCCACCACCACCCAGACCTTCAAACCCACCCCTCTCTTTTAGGA	1078
Db	321	GlyProCysProProProProProArgProSerLysProThrProSerLeuGly	340
QY	1079	CCTGCTGGTGACAAACCCCCTGGAGCTGAGTAGGATTCCCTGATGAAAACTGCCAGATTAAAC	1138
Db	341	ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn	360
QY	1139	AGATATGGCACTTTCAAGCCACCATCAAAATTGTAGAGGGAATCTTAGAGTTAACATC	1198
Db	361	ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle	380
QY	1199	ATCCAGATGACAGACGTCTGTATGCCGTGCCATGGCCTGAAAGCTCCCTAATAGACTTT	1258
Db	381	IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeulleAspPhe	400
QY	1259	GTCGTGACCTGCCAAGGGAGCATTCCCCACGGAGGTCTGTACCATCATTTCTGACCCCAACC	1318
Db	401	ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr	420
QY	1319	TGCGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTCTG	1378
Db	421	CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu	440
QY	1379	ACTGTGAGACGAACTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGGAT	1438
Db	441	ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp	460
QY	1439	GACACAAGCCTGGCTCTCACGAGCACCCCTGATTCTGTCTTCTGACAGAGACCCAGCCTCG	1498
Db	461	AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer	480
QY	1499	CCTTTAAGGATGGCAAAACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACT	1558
Db	481	ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr	500
QY	1559	GTGATCTCCCTCTTGGTGTACAAAAACACAAGGAATACAACCCCAATAGAAAATAGTCCT	1618
Db	501	ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro	520
QY	1619	GGGAATGTGTCAGAAAGCAAGGCCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCT	1678
Db	521	GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe	540
QY	1679	TTCCCGGGAACCCAGGAAAAGGATCCGCTACTCAAAAACCAAGAATTTAAAGGAGTTCT	1738
Db	541	PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer	560

RESULT 8

US-10-117-982-225
; Sequence 225, Application US/10117982
; Publication No. US20030138438A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Mericle, Barbara

; APPLICANT: Spies, Gregory A.
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C18
; CURRENT APPLICATION NUMBER: US/10/117,982
; CURRENT FILING DATE: 2002-04-05
; NUMBER OF SEQ ID NOS: 484
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-117-982-225

Alignment Scores:
Pred. No.: 7.23e-270 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 14 Gaps: 0

US-10-039-272-1 (1-2661) x US-10-117-982-225 (1-560)

QY	60	ATGGAATGTC	TACTATTCTCTGGGATTTCTGCTCCTGGTGC	AAGATTGCCACTTGAT	119
Db	1	MetGluCysLeuTyr	TyrPheLeuGlyPheLeuLeuAlaA	ArgLeuProLeuAsp	20
QY	120	GCGGCCAA	ACGATTTTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG	179	
Db	21	AlaAlaLysArgPheHis	AspValLeuGlyAsnGluArgProSer	AlaTyrMetArgGlu	40
QY	180	CACAATCA	ATTAATGGCTGGTCTTCTGATGAAATGACTGGAATGAA	AAACTCTACCCA	239
Db	41	HisAsnGlnLeu	AsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro	60	
QY	240	GTGTGGA	AGCGGGAGACATGAGGTGGAAAACTCTTGGAAAGGAGGCCGTGTGCAGCGG	299	
Db	61	ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla	80		
QY	300	GTCCTG	ACCAGTGACTCACCAGCCCTCGTGGGCTCAAATATACATTTGCGGTGAACCTG	359	
Db	81	ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu	100		
QY	360	ATATTCC	TAGATGCCAAAGGAAGATGCCAATGCCAATAGTCTATGAGAGAACTGC	419	
Db	101	IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys	120		
QY	420	AGAAATG	AGGCTGTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG	479	
Db	121	ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu	140		
QY	480	GACAGTG	ACGGGAAAATGGCACCGCCAAAGCCATCATACGTCTTCCCTGATGGGAA	539	
Db	141	AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys	160		
QY	540	CCTTTTC	CTCACACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCTTCCACACACTT	599	
Db	161	ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu	180		
QY	600	GGTCAG	TATTTCCAGAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT	659	
Db	181	GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn	200		
QY	660	GTGACAC	TTGGGCCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATAT	719	
Db	201	ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr	220		
QY	720	GTTCCCA	TCCGACAAGTGAAGATGTGACGTGGTAAACAGATCAGATTCTCTGTGTTGTG	779	
Db	221	ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal	240		

QY	780	ACTATG	TTCAGAAAGACGATCGAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC	838
Db	241	ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro	260	
QY	839	ATTATG	TTTGATGTCCTGATTTCATGATCCTAGCCACTTCTCAATTAATCTACCATTAAAC	898
Db	261	IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn	280	
QY	899	TACAA	TGGAGCTTCGGGATAATACTGGCCTGTTGTTTCCACCAATCATCTGTGAAT	958
Db	281	TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn	300	
QY	959	CACAC	GTATGTCTCAATGGAACTTCAGCCTTAAACCTCACTGTGAAAGCTGCAGCACCA	1018
Db	301	HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro	320	
QY	1019	GGACCT	TGTCGCCACCGCCACCAACCCAGACCTTCAAAACCCACCCCTTCTTTAGGA	1078
Db	321	GlyProCysProProProProProProProProProProProProProProProSerLeuGly	340	
QY	1079	CCTGCT	GGTGACAAACCCCTGGAGCTGAGTAGGATTCTTGATGAAAACTGCCAGATTAAAC	1138
Db	341	ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn	360	
QY	1139	AGATAT	GGCCACTTTCAAGCCACCATCACAAATTGTAGAGGGAATCTTAGAGGTTAAACATC	1198
Db	361	ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle	380	
QY	1199	ATCCAG	ATGACAGACGCTCCTGATGCCGGTGCCATGGCTGAAAGCTCCCTAATAGACTTT	1258
Db	381	IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe	400	
QY	1259	GTGCTG	ACCTGCCAAGGGAGCATTCCCACGGAGGTCTGTACCATCATTTCTGACCCACC	1318
Db	401	ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr	420	
QY	1319	TGCGAG	ATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGCTG	1378
Db	421	CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu	440	
QY	1379	ACTGTG	AGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGAT	1438
Db	441	ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp	460	
QY	1439	GACACA	AGCCTGCTCACGAGCACCCCTGATTTCTGTTCCTGACAGACACCCAGCCTCG	1498
Db	461	AspThrSerSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer	480	
QY	1499	CCTTTA	AGGATGGCAAAACAGTCCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACT	1558
Db	481	ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr	500	
QY	1559	GTGATC	TCTCCCTCTTGGTGTAACAAAAACACAAGGAATACAACCCCAATAGAAAAATAGTCCT	1618
Db	501	ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro	520	
QY	1619	GGGAAT	GTGGTCAGAAAGGCTGAGTGTCTTTCTCAACCCGTGCAAAAAGCCGTGTTT	1678
Db	521	GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe	540	
QY	1679	TTCCCC	GGGAAACACGAGAAAGGATCCGCTACTCAAAAAACCAAGAAATTTAAAGGAGTTTCT	1738
Db	541	PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer	560	

RESULT 9
US-10-295-027-1258
; Sequence 1258, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.


```
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hezezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1258
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1258

Alignment Scores:
Pred. No.: 7.23e-270 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservatives: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 15 Gaps: 0

US-10-039-272-1 (1-2661) x US-10-295-027-1258 (1-560)
QY 60 ATGGAATGTCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTTCATGATGTCTGCTGCAATGAAAGACCTTCTGCTTACATGAGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAATGGCTGCTTCTGATGAAAATGACTGGAATGAAAACTCTACCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAAGCGGGGAGACATGAGGTGGAATACTCCTGGAAGGAGCGCGTGTGCAGGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80
QY 300 GTCCTGACCAGTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCTTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
```

```
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGCTGGTTTATCTCTGATCCATATGTTTACAACTGGACAGCATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGAAAAATGGACCGGCCAAAGCCCATCATACGTTCTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACCACCCCGGATGGAGAAAGATGGAATTTTCATCTACGTCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAGTATTTCCAGAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCCTCAACTCATGGAAGTACTGTCTACAGAAGACATGGACGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTGATGTCCTGATTTCATGATCCTTAGCCACTTCCCTCAATTATCTACCATTAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGCGCTGTTTGTTCACCAATCATACTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTGCTCAATGGAACCTTCAGCCCTTAACCTCACTGTGAAAGCTGCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCCACCACCCAGACCTTCAAAACCCACCTTCTTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProProProProProProSerLeuGly 340
QY 1079 CCTGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACCTGCCAGATTAAAC 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTTCAAGCCACCACCATCACAAATGTAGAGGGAATCTTAGAGGTTAACATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGTCCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACCTGCCAAGGGAGCATTTCCCACGGAGGTCGTGTACCATCATTTTCTGACCCACC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TGGGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGTGTCTGTCTGTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACACAAGCCTGGCTCTCACGAGCACCCCTGATTTCTGTTCTCTGACAGAGACCCAGCCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
```

QY 1499 CCTTTAAGGATGCAAAACAGTCCCTGATCTCCGTTGGCTGGCTGGCCATATTTGTCACT 1558
Db |||||||
481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTTGGTGACAAAAACACAAAGGAATACAAACCCATAGAAAAATAGTCCT 1618
Db |||||||
501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGAATGTGGTCAGAAAGCAAGGCGCTAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTTC 1678
Db |||||||
521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGGAACAGGAAAGGATCCGCTACTCAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db |||||||
541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 10

US-10-313-986-225
; Sequence 225, Application US/10313986
; Publication No. US20030236209A1
; GENERAL INFORMATION:
; APPLICANT: Foy, Teresa M.
; APPLICANT: McNabb, Andria
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Reed, Steven G.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C19
; CURRENT APPLICATION NUMBER: US/10/313,986
; CURRENT FILING DATE: 2002-12-04
; NUMBER OF SEQ ID NOS: 560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 225
; LENGTH: 560
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-313-986-225

Alignment Scores:
Pred. No.: 7.23e-270 Length: 560
Score: 2999.00 Matches: 559
Percent Similarity: 99.82% Conservative: 0
Best Local Similarity: 99.82% Mismatches: 1
Query Match: 61.84% Indels: 1
DB: 15 Gaps: 0

US-10-039-272-1 (1-2661) x US-10-313-986-225 (1-560)

QY 60 ATGGAATGTCTCTACTATTTCCTGGGATTTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db |||||||
1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTCATGATGTCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db |||||||
21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAAATGGCTGGTCTTCTGATGATAAATGACTGGAATGAAAACTCTACCCA 239
Db |||||||
41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAAGCGGGGACATGAGGTGGAAAACTCCTGGAAGGAGGCCGTGTGCAGGCG 299
Db |||||||
61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80
QY 300 GTCCTGACCAGTGACTCACCAGCCCTCGTGGGTCAAATATACATTGGCGTGAACCTG 359
Db |||||||
81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db |||||||
101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120

QY 420 AGAAATGAGGTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGGTCAGAG 479
Db |||||||
121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGAAAAATGGCACCCGCCAAAGCCCATATAACGTCCTTCCCTGATGGGAAA 539
Db |||||||
141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACCAACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCTTCCACACACTT 599
Db |||||||
161 PropheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAGTATTTCCAGAAATTTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db |||||||
181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCCTCAACTCATGGAAGTGACTGTCTACAGAAACATGGACGGGCATAT 719
Db |||||||
201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCCTGTGTTGTG 779
Db |||||||
221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db |||||||
241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTGATGTCCTGATTCATGATCCTAGCCACTTCCCTCAATTATTCTACCATTAAC 898
Db |||||||
261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAAGTGGAGCTTCGGGGATAATACTAGCCCTGTTTGTTCACCAATCATACTGTGAAAT 958
Db |||||||
281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCTTAACTCACTGTGAAAGCTGCAGCACCA 1018
Db |||||||
301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCACCCAGACCTTCAAAACCCACCCCTTCTTTAGGA 1078
Db |||||||
321 GlyProCysProProProProProProProProProProProProProProProSerLeuGly 340
QY 1079 CCTGCTGTTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACCTGCCAGATTAAAC 1138
Db |||||||
341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTTCAAGCCACCACCATCACAAATGTAGAGGGAATCTTAGAGGTAAACATC 1198
Db |||||||
361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACGTCCTGATGCCGGTGCCATGGCCTGAAAGCTCCCTAATAGACTTT 1258
Db |||||||
381 IleGlnMetThrAspValLeuMetProValProTrpProGluSerSerLeuIleAspPhe 400
QY 1259 GTCGTGACCTGCCAAGGGAGCATTCACAGGAGGTCTGTACCATCATTTCTGACCCACC 1318
Db |||||||
401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TCGGAGATCACCCAGAAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTCTG 1378
Db |||||||
421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCCCTGGGGAT 1438
Db |||||||
441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACACAAGCCTGGCTCTCACGAGCACCCCTGATTTTCTGTTCCCTGCAGAGACCCAGCCTCG 1498
Db |||||||
461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480

Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGATATACTAGCCCTGTTTGTTCACCAATCATACTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCAGTGTGAAAGCTGCAGCACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCACCCAGACCTTCAAAACCCACCCCTTCTTTAGGA 1078
Db 321 GlyProCysProProProProProProProProProProProProProProSerLeuGly 340
QY 1079 CCTGCTGGTGACAAACCCCTGGAGCTGAGTGGATTCCTGATGAAAGCTGCCAGATTAAC 1138
Db 341 ProAlaGlyAspAsnProLeuGluLeuSerArgIleProAspGluAsnCysGlnIleAsn 360
QY 1139 AGATATGGCCACTTCAAGCCACCACCATCACAAATTGTAGAGGAATCTTAGAGGTTAAACATC 1198
Db 361 ArgTyrGlyHisPheGlnAlaThrIleThrIleValGluGlyIleLeuGluValAsnIle 380
QY 1199 ATCCAGATGACAGACCTCTGATGCCGGTGCATGGCCTGAAAGCTCCCTAAATAGACTTT 1258
Db 381 IleGlnMetThrAspValLeuMetProValProProProProProProProProProPhe 400
QY 1259 GTCGTGACCTGCCAAGGAGCATTCCCACGGAGTCTGTACCATCATTTCTGACCCACC 1318
Db 401 ValValThrCysGlnGlySerIleProThrGluValCysThrIleIleSerAspProThr 420
QY 1319 TGCAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTG 1378
Db 421 CysGluIleThrGlnAsnThrValCysSerProValAspValAspGluMetCysLeuLeu 440
QY 1379 ACTGTGAGACGAACCTTCAATGGTCTGGGACGTACTGTGTGAACCTCACCTGGGGGAT 1438
Db 441 ThrValArgArgThrPheAsnGlySerGlyThrTyrCysValAsnLeuThrLeuGlyAsp 460
QY 1439 GACACAAAGCTGCTCTCAGGACACCTGATTTCTGTCTCTGACAGAGACCCAGCTCG 1498
Db 461 AspThrSerLeuAlaLeuThrSerThrLeuIleSerValProAspArgAspProAlaSer 480
QY 1499 CCTTTAAGGATGGCAACACAGTCCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACT 1558
Db 481 ProLeuArgMetAlaAsnSerAlaLeuIleSerValGlyCysLeuAlaIlePheValThr 500
QY 1559 GTGATCTCCCTCTTGGTGTACAAAAACACAAGGAATACACCCCAATAGAAAATAGTCT 1618
Db 501 ValIleSerLeuLeuValTyrLysLysHisLysGluTyrAsnProIleGluAsnSerPro 520
QY 1619 GGGATGTGGTCAGAAAGCAAGGCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTC 1678
Db 521 GlyAsnValValArgSerLysGlyLeuSerValPheLeuAsnArgAlaLysAlaValPhe 540
QY 1679 TTCCCGGAAACCAAGGATCCGCTACTCAAAACCCAGAATTTAAAGGAGTTTCT 1738
Db 541 PheProGlyAsnGlnGluLysAspProLeuLeuLysAsnGlnGluPheLysGlyValSer 560

RESULT 13
US-10-219-535-42
; Sequence 42, Application US/10219535
; Publication No. US2004004179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.

; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3530P1C60
; CURRENT APPLICATION NUMBER: US/10/219,535
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 42
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-219-535-42

Alignment Scores:
Pred. No.: 5,32e-268 Length: 572
Score: 2979.00 Matches: 558
Percent Similarity: 97.55% Conservative: 0
Best Local Similarity: 97.55% Mismatches: 2
Query Match: 61.42% Indels: 13
DB: 12 Gaps: 1

US-10-039-272-1 (1-2661) x US-10-219-535-42 (1-572)

QY 60 ATGGAATGTCTCTACTATTCTCTGGGATTTCTGCTGGCTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTTCATGATGTCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAAATGGCTGCTTCTGTGATGAAAATGACTGGAATGAAAACTTACCCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGAAAGCGGGAGACATGAGGTGAAAAAACTCCTGGAAGGAGGCCGTGTGAGGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyArgValGlnAla 80
QY 300 GTCCTGACCAAGTCACTCACCGCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGCTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGAAAATGGCACCGGCCAAAGCCATCATACGTCTTCCTGATGGGAAA 539

Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACCACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTyrArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAGTATTTCCAGAAATTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCCTCAACTCATGGAAGTGACTGTCTACAGAGACATGGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAACAGATCAGATTCTCTGTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAGAACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTGATGTCCTGATTCATGATCCTAGCCACTTCTCAATTAATTTACCAATTAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATACTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGAACCTTCAGCCTTAACCTCACTGTGAAGCTGCAGACCA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCACCCACCCAGACCTTCAAAACCCACCTTCTTA--- 1075
Db 321 GlyProCysProProProProProProProProProProProProProProProProPro 340
QY 1076 -----GGACCTGCTGGTGACAAACCCCTGGAG 1102
Db 341 ThrThrLeuLysSerTyrAspSerAsnThrProGlyProThrGlyAspAsnProLeuGlu 360
QY 1103 CTGAGTAGGATTCCTGATGAAACTGCCAGATTAAACAGATATGGCCACTTCAAGCCACC 1162
Db 361 LeuSerArgIleProAspGluAsnCysGlnIleAsnArgTyrGlyHisPheGlnAlaThr 380
QY 1163 ATCACAATTGTAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACCTCCTGATG 1222
Db 381 IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet 400
QY 1223 CCGGTGCCATGGCTGAAAGCTCCCTAATAGACTTTGTCGTGACCTGCCAAGGGAGCATT 1282
Db 401 ProValProTrpProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle 420
QY 1283 CCCACGGAGGTCTGTACCATCATTTCTGACCCCACTGCGAGATCACCAGAACACAGTC 1342
Db 421 ProThrGluValCysThrIleIleSerAspProThrCysGluIleThrGlnAsnThrVal 440
QY 1343 TGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTCTGACTGTGAGACGACCTTCAATGGG 1402
Db 441 CysSerProValAspValAspGluMetCysLeuLeuThrValArgArgThrPheAsnGly 460
QY 1403 TCTGGGACGTACTGTGTGAACCTCACCTGGGGGATGACACAAGCCTGGCTCTCACGAGC 1462
Db 461 SerGlyThrTyrCysValAsnLeuThrLeuGlyAspThrSerLeuAlaLeuThrSer 480
QY 1463 ACCCTGATTTCTGTTCTTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAACAGTGCC 1522
Db 481 ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla 500
QY 1523 CTGATCTCCGTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAA 1582

Db 501 LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuLeuValTyrLys 520
QY 1583 AAACACAAGGAATACAACCCCAATAGAAAAATAGTCTGGGAATGTGGTCAAGCAAGGC 1642
Db 521 LysHisLysGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerLysGly 540
QY 1643 CTGAGTGTCTTTCTCAACCCGTGCAAAAGCCGTGTCTTCCCGGGAACCCAGGAAAGGAT 1702
Db 541 LeuSerValPheLeuAsnArgAlaLysAlaValPhePheProGlyAsnGlnGluLysAsp 560
QY 1703 CCGTACTCAAAACCAAGAATTAAAGGAGTTTCT 1738
Db 561 ProIleLeuLysAsnGlnGluPheLysGlyValSer 572

RESULT 14

US-10-232-230-42
; Sequence 42, Application US/10232230
; Publication No. US2004004180A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C103
; CURRENT APPLICATION NUMBER: US/10/232,230
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 42
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-232-230-42

Alignment Scores:
Pred. No.: 5.32e-268 Length: 572
Score: 2979.00 Matches: 558
Percent Similarity: 97.55% Conservative: 0
Best Local Similarity: 97.55% Mismatches: 2
Query Match: 61.42% Indels: 13
DB: 12 Gaps: 1

US-10-039-272-1 (1-2661) x US-10-232-230-42 (1-572)

QY 60 ATGGAATGCTCTACTATTCTCTGGGATTCTCTGCTCTGGTGCAAGATTGCCACTTGAT 119
Db 1 MetGluCysLeuTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20
QY 120 GCCGCCAAACGATTTCATGATGTCTGGCAATGAAAGACCTTCTGCTTACATGAGGGAG 179
Db 21 AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40
QY 180 CACAATCAATTAATGGCTGCTTCTGATGAAAAATGACTGGAATGAAAACTCTACCCCA 239
Db 41 HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60
QY 240 GTGTGGAGCGGGAGACATGAGGTGGAAAAAATCCTGGAGGAGGCGCGTGTGCAGGCG 299
Db 61 ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80
QY 300 GTCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTG 359
Db 81 ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100
QY 360 ATATTCCCTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db 101 IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120
QY 420 AGAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG 479
Db 121 ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140
QY 480 GACAGTGACGGGAAAAATGGCACCGGCCAACCCATCATAACGCTCTTCCCTGATGGGAAA 539
Db 141 AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160
QY 540 CCTTTTCTCACCACCCCGGATGGAGAGATGGAATTTCACTACGCTTCCACACACTT 599
Db 161 ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180
QY 600 GGTCAAGTATTTCCAGAAATTTGGACGAGTTCAGTGAGAGTTTCTGTGAACACAGCCCAAT 659
Db 181 GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200
QY 660 GTGACACTTGGGCTCAACTCATGGAAGTACTGTCTACAGAAAGACATGGACGGGCATAT 719
Db 201 ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
QY 720 GTTCCCATCGCACAGTGAAAGATGTGTACGTGGTGAACAGATCAGATTCTCTGTGTTGTG 779
Db 221 ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240
QY 780 ACTATGTTCCAGAAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db 241 ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260
QY 839 ATTATGTTTGATGTCCTGATTCATGATCCTAGCCACTTCTCAATATTATTCTACCAATAAC 898
Db 261 IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280
QY 899 TACAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATCTGTGAAT 958
Db 281 TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300
QY 959 CACACGTATGTCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCACA 1018
Db 301 HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320
QY 1019 GGACCTTGTCCGCCACCGCCACCACCCAGACCTTCAAACCCACCCCTTCTTTA--- 1075
Db 321 GlyProCysProProProProProProProProProProProProProProProProPro 340
QY 1076 -----GGACCTGCTGGTGACAAACCCCTGGAG 1102
Db 341 ThrThrLeuLysSerTyrAspSerAsnThrProGlyProThrGlyAspAsnProLeuGlu 360
QY 1103 CTGAGTAGGATTCCTGATGAAAACTGCCAGATTAACAGATATGGCCACTTTCAAGCCACC 1162

Db 361 LeuSerArgIleProAspGluAsnCysGlnIleAsnArgTyrGlyHisPheGlnAlaThr 380
QY 1163 ATCACAATTTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCCTGATG 1222
Db 381 IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet 400
QY 1223 CCGTGCCCATGGCCTGAAAGCTCCCTAAATAGACTTTTGTGCTGACCTGCCAAGGGAGCAT 1282
Db 401 ProValProTrpProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle 420
QY 1283 CCCACGGAGGTCTGTACCATCATTTCTGACCCACCTGGAGATCACCCAGAACACAGTC 1342
Db 421 ProThrGluValCysThrIleIleSerAspProThrCysGluIleThrGlnAsnThrVal 440
QY 1343 TGCAGCCTGTGGATGTGGATGAGATGTGTCTGTGCTGACTGTGAGACGAACCTTCAATGGG 1402
Db 441 CysSerProValAspValAspGluMetCysLeuLeuThrValArgArgThrPheAsnGly 460
QY 1403 TCTGGGACGTACTGTGTGAACCTCACCTGGGGGATGACACAAGCCTGGCTCTCACGAGC 1462
Db 461 SerGlyThrTyrCysValAsnLeuThrLeuGlyAspThrSerLeuAlaLeuThrSer 480
QY 1463 ACCCTGATTCTCTCTGACAGAGACCCAGCCTCGCCTTTAAGGATGCGCAAAACAGTGCC 1522
Db 481 ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla 500
QY 1523 CTGATCTCCGTTGGCTGCTGGCCATATTGTCACTGTGATCTCCCTTGGTGTACAAA 1582
Db 501 LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuLeuValTyrLys 520
QY 1583 AAACACAAGGAATACAAACCAATAGAAAATAGTCTGGGAATGTGTCAGAAAGCAAGGC 1642
Db 521 LysHisLysGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerLysGly 540
QY 1643 CTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTTCTTCCCGGAAACCCAGGAAAAGGAT 1702
Db 541 LeuSerValPheLeuAsnArgAlaLysAlaValPhePheProGlyAsnGlnGluLysAsp 560
QY 1703 CCGCTACTCAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db 561 ProLeuLeuLysAsnGlnGluPheLysGlyValSer 572

RESULT 15

US-10-232-224-42
; Sequence 42, Application US/10232224
; Publication No. US20030065147A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3530P1C111
; CURRENT APPLICATION NUMBER: US/10/232,224
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103

```
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 42
; LENGTH: 572
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-232-224-42

Alignment Scores:
Pred. No.:      5.32e-268      Length:      572
Score:          2979.00      Matches:      558
Percent Similarity: 97.55%      Conservative: 0
Best Local Similarity: 97.55%      Mismatches:  2
Query Match:     61.42%      Indels:      13
DB:              12          Gaps:         1

US-10-039-272-1 (1-2661) x US-10-232-224-42 (1-572)

QY      60  ATGGAATGTCCTACTATTTCCTGGGATTCTCTGCTCCTGGCTGCAAGATTGCCACTTGAT 119
Db      1  MetGluCysLeuTyrTyrPheLeuGlyPheLeuLeuAlaAlaArgLeuProLeuAsp 20

QY      120 GCCGCCAAACGATTTTCATGATGTGTGGGCAANTGAAAGACCTTCTGCTTACATGAGGGAG 179
Db      21  AlaAlaLysArgPheHisAspValLeuGlyAsnGluArgProSerAlaTyrMetArgGlu 40

QY      180 CACAATCAATTAAATGGCTGTCTTCTGATGAAATGACTGGATGAAATAAAGAACTCTACCCA 239
Db      41  HisAsnGlnLeuAsnGlyTrpSerSerAspGluAsnAspTrpAsnGluLysLeuTyrPro 60

QY      240 GTGTGGAAGCGGGGAGACATGAGGTGGAAAAAACTCCTGGAGGGAGCCGTGTGCAGGCG 299
Db      61  ValTrpLysArgGlyAspMetArgTrpLysAsnSerTrpLysGlyGlyArgValGlnAla 80

QY      300 GTCCTGACCAGTGACTCACAGACCCCTCGTGGGCTCAAATATAAACATTTGCGGTGAACCTG 359
Db      81  ValLeuThrSerAspSerProAlaLeuValGlySerAsnIleThrPheAlaValAsnLeu 100

QY      360 ATATTCCCTAGATGCCAAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAAGAACTGC 419
Db      101  IlePheProArgCysGlnLysGluAspAlaAsnGlyAsnIleValTyrGluLysAsnCys 120

QY      420 AGAAATGAGGCTGGTTTATCTGTGATCCATATGTTTACAACCTGGACAGCATGGTCAGAG 479
Db      121  ArgAsnGluAlaGlyLeuSerAlaAspProTyrValTyrAsnTrpThrAlaTrpSerGlu 140

QY      480 GACAGTGACGGGAAAATGGCACCGGCCAAAGCCCATCATAAAGTCTTCCCTGATGGGAAA 539
Db      141  AspSerAspGlyGluAsnGlyThrGlyGlnSerHisAsnValPheProAspGlyLys 160

QY      540 CCTTTTCCTCACCACCCCGGATGGAGAAGATGGAATTTTCATCTACGTCTTCCCTGATGGGAAA 599
Db      161  ProPheProHisHisProGlyTrpArgArgTrpAsnPheIleTyrValPheHisThrLeu 180

QY      600 GGTCAGTATTTCCAGAAATTGGGACGATGTTCAAGTGAGAGTTTCTGTGAACACAGCCAAT 659
Db      181  GlyGlnTyrPheGlnLysLeuGlyArgCysSerValArgValSerValAsnThrAlaAsn 200

QY      660 GTGACACTTGGGCCTCAACTCATGGAAGTGACTGTCTACAGAAGACATGGACGGGCATAT 719
Db      201  ValThrLeuGlyProGlnLeuMetGluValThrValTyrArgArgHisGlyArgAlaTyr 220
```

```
QY      720 GTTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAACAGATCAGATTCTCTGTGTTGTG 779
Db      221  ValProIleAlaGlnValLysAspValTyrValValThrAspGlnIleProValPheVal 240

QY      780 ACTATGTTCCAGAAGACGATCGAAATTCATCCGACGAAACCTTCCC-AAAGATCTCCCC 838
Db      241  ThrMetPheGlnLysAsnAspArgAsnSerSerAspGluThrPheLeuLysAspLeuPro 260

QY      839 ATTATGTTTGATGTCCTGATTCATGATCCTAGCCACTTCTCAATATTATTACCAATAAC 898
Db      261  IleMetPheAspValLeuIleHisAspProSerHisPheLeuAsnTyrSerThrIleAsn 280

QY      899 TACAAGTGGAGCTTCGGGGATAATACTGGCCCTGTTGTTCCACCAATCATACTACTGAAT 958
Db      281  TyrLysTrpSerPheGlyAspAsnThrGlyLeuPheValSerThrAsnHisThrValAsn 300

QY      959 CACACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCAGTGTGAAAGCTGCAGACCA 1018
Db      301  HisThrTyrValLeuAsnGlyThrPheSerLeuAsnLeuThrValLysAlaAlaPro 320

QY      1019 GGACCTTGTCCGCCACCGCCACCCACCCAGACCTTCAAAAACCCACCCCTTCTTTA--- 1075
Db      321  GlyProCysProProProProProProProProProProProProProProProProSerLeuAla 340

QY      1076 -----GGACCTGTGCTGTCACAAACCCCTGGAG 1102
Db      341  ThrThrLeuLysSerTyrAspSerAsnThrProGlyProThrGlyAspAsnProLeuGlu 360

QY      1103 CTGAGTAGGATTTCCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCGAAGCCACC 1162
Db      361  LeuSerArgIleProAspGluAsnCysGlnIleAsnArgTyrGlyHisPheGlnAlaThr 380

QY      1163 ATCACAATTGTAGAGGGAATCTTAGAGGTTAATCATCATCCAGATGACAGACGTCTGATG 1222
Db      381  IleThrIleValGluGlyIleLeuGluValAsnIleIleGlnMetThrAspValLeuMet 400

QY      1223 CCGTGCCCATGGCTGAAAGCTCCCTAATAGACTTTTGCTGACCTGCCAAGGGAGCATT 1282
Db      401  ProValProTrpProGluSerSerLeuIleAspPheValValThrCysGlnGlySerIle 420

QY      1283 CCACGCGAGGTCTGTACCATCATTTTCTGACCCCCACCTGCGAGATCACCCAGAACACAGTC 1342
Db      421  ProThrGluValCysThrIleIleSerAspProThrCysGluIleThrGlnAsnThrVal 440

QY      1343 TGCAGCCCTGTGGATGCGATGAGATGTGTCTGTGACTGTGAGACGAAACCTTCAATGGG 1402
Db      441  CysSerProValAspValAspGluMetCysLeuLeuThrValArgArgThrPheAsnGly 460

QY      1403 TCTGGGACGTACTGTGTGAACCTCACCCCTGGGGATGACACAAGCCTGGCTCTCACGAGC 1462
Db      461  SerGlyThrTyrCysValAsnLeuThrLeuGlyAspAspThrSerLeuAlaLeuThrSer 480

QY      1463 ACCCTGATTTCTGTTCCTGACAGAGACCCAGCCCTCGCCTTTAAGGATGGCAAAACAGTGCC 1522
Db      481  ThrLeuIleSerValProAspArgAspProAlaSerProLeuArgMetAlaAsnSerAla 500

QY      1523 CTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACGTGTGATCTCCCTCTTGGTGTAACA 1582
Db      501  LeuIleSerValGlyCysLeuAlaIlePheValThrValIleSerLeuLeuValTyrLys 520

QY      1583 AAACACAAGGAATACACCCCAATAGAAAATAGTCTCTGGGAATGTGTCAGAAGCAAAAGGC 1642
Db      521  LysHisLysGluTyrAsnProIleGluAsnSerProGlyAsnValValArgSerLysGly 540

QY      1643 CTGAGTGTCTTTCTCAACCGTGCAAAAGCCCGTGTCTTCCCGGGAACCCAGGAAAAGGAT 1702
Db      541  LeuSerValPheLeuAsnArgAlaLysAlaValPhePheProGlyAsnGlnGluLysAsp 560

QY      1703 CCGCTACTCAAAAACCAAGAAATTTAAAGGAGTTTCT 1738
Db      561  ProLeuLeuLysAsnGlnGluPheLysGlyValSer 572
```

us-10-039-272-1.rapb

Mon Sep 13 10:28:10 2004

Search completed: September 10, 2004, 16:38:13
Job time : 613.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.
OM nucleic - nucleic search, using sw model
Run on: September 10, 2004, 07:08:49 ; Search time 3913 Seconds
(without alignments)
3421.641 Million cell updates/sec
Title: US-10-039-272-1
Perfect score: 2661
Sequence: 1 cggcacgagggccagagga.....aaaaaaaaaaaaaaaaaaaa 2661
Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0
Searched: 3304383 seqs, 2515761380 residues
Total number of hits satisfying chosen parameters: 6608766
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq2.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				%		Description	
Result No.	Score	Query Match	Length DB ID				
1	2661	100.0	2661 13	US-10-463-106-1		Sequence 1, Appli	
2	2661	100.0	2661 14	US-10-039-272-1		Sequence 1, Appli	
3	2559.2	96.2	2952 15	US-10-198-846-10843		Sequence 10843, A	
4	2547.8	95.7	2669 10	US-09-525-978B-82		Sequence 82, Appl	
5	2547.8	95.7	2669 13	US-10-342-887-779		Sequence 779, App	
6	2547.8	95.7	2669 13	US-10-172-118-779		Sequence 779, App	
7	2547.8	95.7	2669 16	US-10-295-027-1041		Sequence 1041, Ap	
8	2547.8	95.7	2669 17	US-10-450-826-92		Sequence 92, Appl	
9	2531.6	95.1	2666 15	US-10-084-817-169		Sequence 169, App	
10	2485	93.4	2683 13	US-10-219-535-41		Sequence 41, Appl	
11	2485	93.4	2683 13	US-10-232-230-41		Sequence 41, Appl	
12	2485	93.4	2683 13	US-10-232-224-41		Sequence 41, Appl	
13	2485	93.4	2683 15	US-10-227-884-41		Sequence 41, Appl	
14	2485	93.4	2683 15	US-10-230-163-41		Sequence 41, Appl	

15	2485	93.4	2683	15	US-10-230-338-41	Sequence 41, Appl
16	2485	93.4	2683	15	US-10-218-631-41	Sequence 41, Appl
17	2485	93.4	2683	15	US-10-230-414-41	Sequence 41, Appl
18	2485	93.4	2683	15	US-10-216-159A-41	Sequence 41, Appl
19	2485	93.4	2683	15	US-10-218-849-41	Sequence 41, Appl
20	2485	93.4	2683	15	US-10-227-873-41	Sequence 41, Appl
21	2485	93.4	2683	15	US-10-227-883-41	Sequence 41, Appl
22	2485	93.4	2683	15	US-10-219-076-41	Sequence 41, Appl
23	2485	93.4	2683	15	US-10-230-434-41	Sequence 41, Appl
24	2485	93.4	2683	15	US-10-219-003-41	Sequence 41, Appl
25	2485	93.4	2683	15	US-10-219-075-41	Sequence 41, Appl
26	2485	93.4	2683	15	US-10-219-464-41	Sequence 41, Appl
27	2485	93.4	2683	15	US-10-219-466-41	Sequence 41, Appl
28	2485	93.4	2683	15	US-10-219-479-41	Sequence 41, Appl
29	2485	93.4	2683	15	US-10-219-481-41	Sequence 41, Appl
30	2485	93.4	2683	15	US-10-230-260-41	Sequence 41, Appl
31	2485	93.4	2683	15	US-10-232-231-41	Sequence 41, Appl
32	2485	93.4	2683	15	US-10-232-233-41	Sequence 41, Appl
33	2485	93.4	2683	15	US-10-216-165-41	Sequence 41, Appl
34	2485	93.4	2683	15	US-10-218-956-41	Sequence 41, Appl
35	2485	93.4	2683	15	US-10-219-468-41	Sequence 41, Appl
36	2485	93.4	2683	15	US-10-219-478-41	Sequence 41, Appl
37	2485	93.4	2683	15	US-10-219-536-41	Sequence 41, Appl
38	2485	93.4	2683	15	US-10-233-205-41	Sequence 41, Appl
39	2485	93.4	2683	15	US-10-219-072-41	Sequence 41, Appl
40	2485	93.4	2683	15	US-10-219-470-41	Sequence 41, Appl
41	2485	93.4	2683	15	US-10-219-474-41	Sequence 41, Appl
42	2485	93.4	2683	15	US-10-219-524-41	Sequence 41, Appl
43	2485	93.4	2683	15	US-10-219-528-41	Sequence 41, Appl
44	2485	93.4	2683	15	US-10-227-880-41	Sequence 41, Appl
45	2485	93.4	2683	15	US-10-227-881-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1
US-10-463-106-1
; Sequence 1, Application US/10463106
; Publication No. US20030202938A1
; GENERAL INFORMATION:
; APPLICANT: RAMESHWAR, Pranela
; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-1 GENE
; FILE REFERENCE: 267/033 University of Medicine & Dentistry of New Jersey
; CURRENT APPLICATION NUMBER: US/10/463,106
; CURRENT FILING DATE: 2003-06-17
; PRIOR APPLICATION NUMBER: US 10/039,272
; PRIOR FILING DATE: 2001-10-20
; PRIOR APPLICATION NUMBER: US 60/241,881
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2661
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-463-106-1

Query Match				100.0%; Score 2661; DB 13; Length 2661;			
Best Local Similarity				100.0%; Pred. No. 0;			
Matches 2661; Conservative				0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CGGCACGAGGGCCAGAGGAATAAGTTAACCTTGGTGCCTGCGTCGTCGAGATTTCAGCA	60				
Db	1	CGGCACGAGGGCCAGAGGAATAAGTTAACCTTGGTGCCTGCGTCGTCGAGATTTCAGCA	60				
QY	61	TGGAATGTCTCTACTATTTCTCGGATTTCTGCTCCTGGTGAAGATTGCCACTTGATG	120				
Db	61	TGGAATGTCTCTACTATTTCTCGGATTTCTGCTCCTGGTGAAGATTGCCACTTGATG	120				
QY	121	CGGCAAAACGATTTCATGATGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAGC	180				
Db	121	CGGCAAAACGATTTCATGATGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAGC	180				

QY	181	ACAAATCAATTAAATGGCTGCTCTCTGATGAAATGACTGGAATGAAAAA	240	1261	CCTGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACCATCATTTCTGACCCACCTG	1320
Db	181	ACAAATCAATTAAATGGCTGCTCTCTGATGAAATGACTGGAATGAAAAA	240	1261	CCTGACCTGCCAAGGGAGCATTTCCACGGAGGTCTGTACCATCATTTCTGACCCACCTG	1320
QY	241	TGTGAAGCGGGAGACATGAGTGGAAAAA	300	1321	CGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTGAC	1380
Db	241	TGTGAAGCGGGAGACATGAGTGGAAAAA	300	1321	CGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTGTCTGTGAC	1380
QY	301	TCCTGACCACTGACTCACCAGCCCTCGTGGCTCAAAATATAACATTTGCGGTGAACCTGA	360	1381	TGTGAGACGAACCTTCAATGGGTCTGGACGTA	1440
Db	301	TCCTGACCACTGACTCACCAGCCCTCGTGGCTCAAAATATAACATTTGCGGTGAACCTGA	360	1381	TGTGAGACGAACCTTCAATGGGTCTGGACGTA	1440
QY	361	TATTCCTTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAA	420	1441	CACAGCCTGGCTCTCACGAGCACCTGATTTCTGTTCCTGACAGAGACCCAGCCTCGCC	1500
Db	361	TATTCCTTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAA	420	1441	CACAGCCTGGCTCTCACGAGCACCTGATTTCTGTTCCTGACAGAGACCCAGCCTCGCC	1500
QY	421	GAAATGAGGCTGGTTTATCTGCTGATCCCATATGTTTACAACCTGACAGCATGGTCAGAGG	480	1501	TTTAAGGATGGCAACACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACTGT	1560
Db	421	GAAATGAGGCTGGTTTATCTGCTGATCCCATATGTTTACAACCTGACAGCATGGTCAGAGG	480	1501	TTTAAGGATGGCAACACAGTGCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACTGT	1560
QY	481	ACAGTGACGGGGAAAAATGGCACCGGCCAAGGCATATAACGTTTCCCTGATGGAAAC	540	1561	GATCTCCCTCTTGGTGTACAAAAACACAAAGGAATACAAACCAATAGAAAAATAGTCTCTGG	1620
Db	481	ACAGTGACGGGGAAAAATGGCACCGGCCAAGGCATATAACGTTTCCCTGATGGAAAC	540	1561	GATCTCCCTCTTGGTGTACAAAAACACAAAGGAATACAAACCAATAGAAAAATAGTCTCTGG	1620
QY	541	CTTTTCTCCTCACCAACCCCGGATGGAGAAGATGGAATTCATCTACGTTTCCACACACTTG	600	1621	GAATGTGGTCAGAAAGCAAGGCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTTCTT	1680
Db	541	CTTTTCTCCTCACCAACCCCGGATGGAGAAGATGGAATTCATCTACGTTTCCACACACTTG	600	1621	GAATGTGGTCAGAAAGCAAGGCTGAGTGTCTTCTCAACCGTGCAAAAGCCGTGTTCTT	1680
QY	601	GTCAGTATTTCCAGAAATTTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCCAATG	660	1681	CCCGGGAACCCAGAAAAAGGATCCGCTACTCAAAAAACCAAGAAATTTAAAGGAGTTTCTTA	1740
Db	601	GTCAGTATTTCCAGAAATTTGGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCCAATG	660	1681	CCCGGGAACCCAGAAAAAGGATCCGCTACTCAAAAAACCAAGAAATTTAAAGGAGTTTCTTA	1740
QY	661	TGACACTTGGGCTCAACTCATGGAAGTGTACGTGTAAGAGATGTTTCAAGAGACATGGACGGGCATATG	720	1741	AATTTGACCTTGTCTGAAGCTCACTTTTCAGTGCCATTTGATGTGAGATGTGCTGGAG	1800
Db	661	TGACACTTGGGCTCAACTCATGGAAGTGTACGTGTAAGAGATGTTTCAAGAGACATGGACGGGCATATG	720	1741	AATTTGACCTTGTCTGAAGCTCACTTTTCAGTGCCATTTGATGTGAGATGTGCTGGAG	1800
QY	721	TTCCCATCGCACAAAGTGAAGATGTACGTGGTAACAGATCAGATTCCTGTGTTGTGA	780	1801	TGGCTATTAAACCTTTTCTCTAAAGATTTATTTGTTAAATAGATATTGTGTTTGGGAAAG	1860
Db	721	TTCCCATCGCACAAAGTGAAGATGTACGTGGTAACAGATCAGATTCCTGTGTTGTGA	780	1801	TGGCTATTAAACCTTTTCTCTAAAGATTTATTTGTTAAATAGATATTGTGTTTGGGAAAG	1860
QY	781	CTATGTTCCAGAAACGATCGAAATTCATCCGACGAAACCTTCCCAAAGATCTCCCAT	840	1861	TTGAAATTTTATAGTTAAATGTCTTTTAGAGATGGGAGAGGATTTATCTGCAGGC	1920
Db	781	CTATGTTCCAGAAACGATCGAAATTCATCCGACGAAACCTTCCCAAAGATCTCCCAT	840	1861	TTGAAATTTTATAGTTAAATGTCTTTTAGAGATGGGAGAGGATTTATCTGCAGGC	1920
QY	841	TATGTTTGATGTCCTGATTCATGATCCTAGCCACTTCTCCTCAATATTTCTACCAATTA	900	1921	AGCTTCAGCCATGTTGTGAACTGTATAAAAGCACTTAGCAAGGCTTCTTTTCATTATTT	1980
Db	841	TATGTTTGATGTCCTGATTCATGATCCTAGCCACTTCTCCTCAATATTTCTACCAATTA	900	1921	AGCTTCAGCCATGTTGTGAACTGTATAAAAGCACTTAGCAAGGCTTCTTTTCATTATTT	1980
QY	901	CAAGTGGAGCTTCGGGGATAATACTGGCCCTGTTTGTTCACCAATCATACTGTGAATCA	960	1981	TTTATGTTTCACTTATAAAGTCTTAGGTAACCTAGTAGATAGAAACACCTGTGTCCCAGG	2040
Db	901	CAAGTGGAGCTTCGGGGATAATACTGGCCCTGTTTGTTCACCAATCATACTGTGAATCA	960	1981	TTTATGTTTCACTTATAAAGTCTTAGGTAACCTAGTAGATAGAAACACCTGTGTCCCAGG	2040
QY	961	CACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCAC	1020	2041	GTAAGGAGAGAGCTACTATTGATTAGAGCCCTAACCCAGGTTAACTGCAAGAGAGCGG	2100
Db	961	CACGTATGTGCTCAATGGAACCTTCAGCCTTAACCTCACTGTGAAAGCTGCAGCAC	1020	2041	GTAAGGAGAGAGCTACTATTGATTAGAGCCCTAACCCAGGTTAACTGCAAGAGAGCGG	2100
QY	1021	ACCTTGTCCGCCACCGCCACCCAGACCTTCAAAACCCACCCCTTCTTTAGGACC	1080	2101	GATACCTTCAGCTTCCATGTAACTGTATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGA	2160
Db	1021	ACCTTGTCCGCCACCGCCACCCAGACCTTCAAAACCCACCCCTTCTTTAGGACC	1080	2101	GATACCTTCAGCTTCCATGTAACTGTATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGA	2160
QY	1081	TGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTAACAG	1140	2161	TCATGTTCCAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTCCTGATGGAACA	2220
Db	1081	TGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAACTGCCAGATTAACAG	1140	2161	TCATGTTCCAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTCCTGATGGAACA	2220
QY	1141	ATATGGCCACTTCAAGCCACCATCACAAATGTAGAGGGAATCTTAGAGTTAAACATCAT	1200	2221	ATAACAGGCCAAGCCTGTGGTATGATGTGCACACTTGTCTAGACTCAGAAAAAATACTAC	2280
Db	1141	ATATGGCCACTTCAAGCCACCATCACAAATGTAGAGGGAATCTTAGAGTTAAACATCAT	1200	2221	ATAACAGGCCAAGCCTGTGGTATGATGTGCACACTTGTCTAGACTCAGAAAAAATACTAC	2280
QY	1201	CCAGATGACAGACTCCTGATGCGCGTGCATGCGCCTGAAAGCTCCCTTAATAGACTTTGT	1260	2281	TCTCATAAATGGGTGGGAGTATTTTGGTGACAACTTCTTTGCTTGGCTGAGTGAAGGAA	2340
Db	1201	CCAGATGACAGACTCCTGATGCGCGTGCATGCGCCTGAAAGCTCCCTTAATAGACTTTGT	1260	2281	TCTCATAAATGGGTGGGAGTATTTTGGTGACAACTTCTTTGCTTGGCTGAGTGAAGGAA	2340
				2341	TGATATTTCATATATTATTTCCATGGACATTTAGTTAGTGTCTTTTATATACCAGGC	2400

```
Db 2341 TGATATTCAATATTCATTTATTCATGGACATTTAGTTAGTCTTTTATATACCAGGC 2400
QY
Db 2401 ATGATGCTGAGTGACACTCTGTGTATATTTCCAAATTTTGTATAGTCGTCACATAT 2460
Db 2401 ATGATGCTGAGTGACACTCTGTGTATATTTCCAAATTTTGTATAGTCGTCACATAT 2460
QY 2461 TTGAAATCAAAATATTAAGACTTTCCAAAAATTTGGTCCCTGGTTTTTCATGGCAACTTG 2520
Db 2461 TTGAAATCAAAATATTAAGACTTTCCAAAAATTTGGTCCCTGGTTTTTCATGGCAACTTG 2520
QY 2521 ATCAGTAAGGATTTCCCTCTGTTTGGAACTPAAACCATTTACTATATGTTAGACAAGAC 2580
Db 2521 ATCAGTAAGGATTTCCCTCTGTTTGGAACTPAAACCATTTACTATATGTTAGACAAGAC 2580
QY 2581 ATTTTTTTTTTCTCTCCTGAAAAAAATGAGGGAAGACAAAAAAATTTT 2640
Db 2581 ATTTTTTTTTTCTCTCCTGAAAAAAATGAGGGAAGACAAAAAAATTTT 2640
QY 2641 AAAAAAAAAAAAAAAAAAAAAA 2661
Db 2641 AAAAAAAAAAAAAAAAAAAAAA 2661

RESULT 2
US-10-039-272-1
; Sequence 1, Application US/10039272
; Publication No. US20020168653A1
; GENERAL INFORMATION:
; APPLICANT: RAMESHWAR, Pranela
; TITLE OF INVENTION: HEMATOPOIETIC GROWTH FACTOR INDUCIBLE NEUROKININ-TYPE
; FILE REFERENCE: 267/033 University of Medicine & Dentistry of New Jersey
; CURRENT APPLICATION NUMBER: US/10/039,272
; CURRENT FILING DATE: 2001-10-20
; PRIOR APPLICATION NUMBER: US 60/241,881
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2661
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-039-272-1

Query Match 100.0%; Score 2661; DB 14; Length 2661;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2661; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCACGAGGGCCCGAGGAATAAGTTAACCTTGGTGCCTGCGTCGCTGAGAAATTCAGCA 60
Db 1 CGGCACGAGGGCCCGAGGAATAAGTTAACCTTGGTGCCTGCGTCGCTGAGAAATTCAGCA 60

QY 61 TGAATGTCTCTACTATTTCCTGGGATTTCTGCTCGCTGGCTGCAAGATTGCCACTTGATG 120
Db 61 TGAATGTCTCTACTATTTCCTGGGATTTCTGCTCGCTGGCTGCAAGATTGCCACTTGATG 120

QY 121 CCGCCAAACGATTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAGC 180
Db 121 CCGCCAAACGATTCATGATGTGCTGGGCAATGAAAGACCTTCTGCTTACATGAGGGAGC 180

QY 181 ACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTCTACCCAG 240
Db 181 ACAATCAATTAAATGGCTGGTCTTCTGATGAAATGACTGGAATGAAAACTCTACCCAG 240

QY 241 TGTGGAAGCGGGAGACATGAGGTGGAATACTCCTGGAAGGGAGCGCTGTGCAGGCGG 300
Db 241 TGTGGAAGCGGGAGACATGAGGTGGAATACTCCTGGAAGGGAGCGCTGTGCAGGCGG 300

QY 301 TCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTGA 360
Db 301 TCCTGACCACTGACTCACCAGCCCTCGTGGGCTCAAATATAACATTTGCGGTGAACCTGA 360

QY 361 TATTCCTTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAACTGCA 420
```

```
Db 361 TATTCCTTAGATGCCAAAGGAAGATGCCAATGGCAACATAGTCTATGAGAACTGCA 420
QY 421 GAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGGTCAGAGG 480
Db 421 GAAATGAGGCTGGTTTATCTGCTGATCCATATGTTTACAACTGGACAGCATGGTCAGAGG 480
QY 481 ACAGTGACGGGAAAAATGGCACCGGCCAAAGCCCATATAACGTTCTTCCTGATGGAAAC 540
Db 481 ACAGTGACGGGAAAAATGGCACCGGCCAAAGCCCATATAACGTTCTTCCTGATGGAAAC 540
QY 541 CTTTTCTCACCAACCCCGGATGGAGAGATGGAATTTTCATCTACGTTCTCCACACTTG 600
Db 541 CTTTTCTCACCAACCCCGGATGGAGAGATGGAATTTTCATCTACGTTCTCCACACTTG 600
QY 601 GTCAGTATTTCCAGAAATTTGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAATG 660
Db 601 GTCAGTATTTCCAGAAATTTGGACGATGTTTCAGTGAGAGTTTCTGTGAACACAGCCAATG 660
QY 661 TGACACTTGGGCCTCAACTCATGGAAGTACTGTCTACAGAGACATGGACGGGCATATG 720
Db 661 TGACACTTGGGCCTCAACTCATGGAAGTACTGTCTACAGAGACATGGACGGGCATATG 720
QY 721 TTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCTCTGTTGTGA 780
Db 721 TTCCCATCGCACAAAGTGAAGATGTGTACGTGGTAAACAGATCAGATTCTCTGTTGTGA 780
QY 781 CTATGTTCCAGAAACGATCGAAATTCATCCGACGAAACCTTCCCAAAGATCTCCCCAT 840
Db 781 CTATGTTCCAGAAACGATCGAAATTCATCCGACGAAACCTTCCCAAAGATCTCCCCAT 840
QY 841 TATGTTGATGTCTGATTTCATGATCCTAGCCACTTCCCTCAATATTCTACCATTAATA 900
Db 841 TATGTTGATGTCTGATTTCATGATCCTAGCCACTTCCCTCAATATTCTACCATTAATA 900
QY 901 CAAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATACTGTGAATCA 960
Db 901 CAAAGTGGAGCTTCGGGGATAATACTGGCCTGTTTGTTCACCAATCATACTGTGAATCA 960
QY 961 CACGTATGTCTCAATGGAACCTTCAGCCTTAACTCACTGTGAAAAGCTGCAGACCAGG 1020
Db 961 CACGTATGTCTCAATGGAACCTTCAGCCTTAACTCACTGTGAAAAGCTGCAGACCAGG 1020
QY 1021 ACCTTGTCGCCACCGCCACCAACCCAGACCTTCAAAACCCACCCCTTCTTTAGGACC 1080
Db 1021 ACCTTGTCGCCACCGCCACCAACCCAGACCTTCAAAACCCACCCCTTCTTTAGGACC 1080
QY 1081 TGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACCTGCCAGATTACAG 1140
Db 1081 TGCTGGTGACAAACCCCTGGAGCTGAGTAGGATTCCTGATGAAAACCTGCCAGATTACAG 1140
QY 1141 ATATGGCCACTTCAAGCCACCATCACAATTTAGAGGGAATCTTAGAGGTTAAACATCAT 1200
Db 1141 ATATGGCCACTTCAAGCCACCATCACAATTTAGAGGGAATCTTAGAGGTTAAACATCAT 1200
QY 1201 CCAGATGACAGACGTCCTGATGCCGGTGCCATGGCCTGAAAAGCTCCCTAATAGACTTTGT 1260
Db 1201 CCAGATGACAGACGTCCTGATGCCGGTGCCATGGCCTGAAAAGCTCCCTAATAGACTTTGT 1260
QY 1261 CGTGACCTGCAAGGGAGCATTCACCGGAGGTCTGTACCATCATTTCTGACCCCCACCTG 1320
Db 1261 CGTGACCTGCAAGGGAGCATTCACCGGAGGTCTGTACCATCATTTCTGACCCCCACCTG 1320
QY 1321 CGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTCTGCTGAC 1380
Db 1321 CGAGATCACCCAGAACACAGTCTGCAGCCCTGTGGATGTGGATGAGATGTCTGCTGAC 1380
QY 1381 TGTGAGACGAACCTTCAATGGTCTGGGACGTAAGTGTGTGAACCTCACCTGGGGATGA 1440
Db 1381 TGTGAGACGAACCTTCAATGGTCTGGGACGTAAGTGTGTGAACCTCACCTGGGGATGA 1440
QY 1441 CACAGCCTGGCTCTCACGAGACCCCTGATTTCTGTTCTTGACAGAGACCCAGCCTCGCC 1500
```


Db 1441 CACAAGCCTGGCTCTCACGAGCACCCCTGATTCTCTGTTCTTGACAGAGACCCAGCCTCGCC 1500

Qy 1501 TTTAAGGATGGCAAAACAGTCCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACTGT 1560

Db 1501 TTTAAGGATGGCAAAACAGTCCCTGATCTCCGTTGGCTGCTTGGCCATATTTGTCACTGT 1560

Qy 1561 GATCTCCCTCTTGGTGACAAAACACAAGGAATACAACCCAAATAGAAATAGTCCCTGG 1620

Db 1561 GATCTCCCTCTTGGTGACAAAACACAAGGAATACAACCCAAATAGAAATAGTCCCTGG 1620

Qy 1621 GAATGTGTCAGAAAGCAAGGCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTT 1680

Db 1621 GAATGTGTCAGAAAGCAAGGCTGAGTGTCTTTCTCAACCGTGCAAAAGCCGTGTTCTT 1680

Qy 1681 CCCGGGAACCCAGGAAAGGATCCGCTACTCAAAAACCAAGAAATTAAGGAGTTCCTTA 1740

Db 1681 CCCGGGAACCCAGGAAAGGATCCGCTACTCAAAAACCAAGAAATTAAGGAGTTCCTTA 1740

Qy 1741 AATTTCGACCTTGTCTGAAGCTCACTTTTCAGTGCCATTGTATGAGATGTCTGGAG 1800

Db 1741 AATTTCGACCTTGTCTGAAGCTCACTTTTCAGTGCCATTGTATGAGATGTCTGGAG 1800

Qy 1801 TGGCTATTAAACCTTTTTCCTAAAGATTATGTTAAATAGATATTGTTGGGAAG 1860

Db 1801 TGGCTATTAAACCTTTTTCCTAAAGATTATGTTAAATAGATATTGTTGGGAAG 1860

Qy 1861 TTGAATTTTATAGGTTAAATGTCAATTTTAGAGATGGGAGAGGATTATACTGCAGGC 1920

Db 1861 TTGAATTTTATAGGTTAAATGTCAATTTTAGAGATGGGAGAGGATTATACTGCAGGC 1920

Qy 1921 AGCTTCAGCCATGTTTGAAACTGATAAAAGCACTTAGCAAGGCTTCTTTTCATTATT 1980

Db 1921 AGCTTCAGCCATGTTTGAAACTGATAAAAGCACTTAGCAAGGCTTCTTTTCATTATT 1980

Qy 1981 TTTATGTTTCACTTATAAAGTCTTAGGTAACTAGTAGATAGAAACACACTGTGCCGAGA 2040

Db 1981 TTTATGTTTCACTTATAAAGTCTTAGGTAACTAGTAGATAGAAACACACTGTGCCGAGA 2040

Qy 2041 GTAAGGAGAGAAGCTACTATTGATTAGAGCCTAACCCAGGTTAACTGCAAGAGGCGG 2100

Db 2041 GTAAGGAGAGAAGCTACTATTGATTAGAGCCTAACCCAGGTTAACTGCAAGAGGCGG 2100

Qy 2101 GATACTTTTCACTTATAAAGTCTTAGGTAACTAGTAGATAGAAACACACTGTGCCGAGA 2160

Db 2101 GATACTTTTCACTTATAAAGTCTTAGGTAACTAGTAGATAGAAACACACTGTGCCGAGA 2160

Qy 2161 TCATGTTCCAAAGCTAACTGAATCCCACTTCAATACACACTCATGAATCCTGTATGGAACA 2220

Db 2161 TCATGTTCCAAAGCTAACTGAATCCCACTTCAATACACACTCATGAATCCTGTATGGAACA 2220

Qy 2221 ATAACAGSCCCAAAGCCTGTGGTATGATGTGCACACTTGCTAGACTCAGAAAAATACTAC 2280

Db 2221 ATAACAGSCCCAAAGCCTGTGGTATGATGTGCACACTTGCTAGACTCAGAAAAATACTAC 2280

Qy 2281 TCTCATAAATGGGTGGAGTATTTTGGTGACAAACCTACTTTTGGCTTGGCTGAGTGAAGGAA 2340

Db 2281 TCTCATAAATGGGTGGAGTATTTTGGTGACAAACCTACTTTTGGCTTGGCTGAGTGAAGGAA 2340

Qy 2341 TGATATTATATATTTTATTTAGTATTTAGTATTTAGTATTTATATATATATATATATATAT 2400

Db 2341 TGATATTATATATTTTATTTAGTATTTAGTATTTAGTATTTATATATATATATATATATAT 2400

Qy 2401 ATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTATAGTCTGTCACATAT 2460

Db 2401 ATGATGCTGAGTGACACTCTTGTGTATATTTCCAAATTTTGTATAGTCTGTCACATAT 2460

Qy 2461 TTGAAATCAAAATATTAAGACTTTCCAAAAATTTGGTCCCTGTTTTCATGGCAACTTG 2520

Db 2461 TTGAAATCAAAATATTAAGACTTTCCAAAAATTTGGTCCCTGTTTTCATGGCAACTTG 2520

Qy 2521 ATCAGTAAGGATTTCCCTCTGTTTGGAACTAAACCACTTACTATATGTTAGACAAGAC 2580

Db 2521 ATCAGTAAGGATTTCCCTCTGTTTGGAACTAAACCACTTACTATATGTTAGACAAGAC 2580

Qy 2581 ATTTTTCCTTCTGAAAAAATGAGGAGACAAAAAATGAGGAGACAAAAAATGAGGAGACAAAAA 2640

Db 2581 ATTTTTCCTTCTGAAAAAATGAGGAGACAAAAAATGAGGAGACAAAAAATGAGGAGACAAAAA 2640

Qy 2641 AAAAAAAAAAAAAAAAAA 2661

Db 2641 AAAAAAAAAAAAAAAAAA 2661

RESULT 3

US-10-198-846-10843

; Sequence 10843, Application US/10198846

; Publication No. US2003009974A1

; GENERAL INFORMATION:

; APPLICANT: Lillie, James

; APPLICANT: Xu, Yongyao

; APPLICANT: Wang, Youzhen

; APPLICANT: Steinmann, Kathleen

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS

; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; TITLE OF INVENTION: THERAPY OF BREAST CANCER

; FILE REFERENCE: MRI-049

; CURRENT APPLICATION NUMBER: US/10/198,846

; CURRENT FILING DATE: 2002-07-18

; PRIOR APPLICATION NUMBER: 60/306,220

; PRIOR FILING DATE: 2001-07-18

; NUMBER OF SEQ ID NOS: 14084

; SOFTWARE: FastSEQ for Windows Version 4.0

; SEQ ID NO 10843

; LENGTH: 2952

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 2950, 2951, 2952

; OTHER INFORMATION: n = A,T,C or G

US-10-198-846-10843

Query Match 96.2%; Score 2559.2; DB 15; Length 2952;

Best Local Similarity 99.0%; Pred. No. 0;

Matches 2606; Conservative 0; Mismatches 23; Indels 3; Gaps 3;

Qy 28 AACCTTGGTGCCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCTCTGGAT 87

Db 187 ACCTTGAGTGCCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTCTACTATTTCTCTGGAT 246

Qy 88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTTCATGATGTGCTGG 147

Db 247 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTTCATGATGTGCTGG 306

Qy 148 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGTGGTCTTCTG 207

Db 307 GCAATGAAAGACCTTCTGCTTACATGAGGGAGCACAAATCAATTAATGGTGGTCTTCTG 366

Qy 208 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGGA 267

Db 367 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGGA 426

Qy 268 AAAACTCCTGGAAGGGAGGCCGTGTGAGGGCGGTCTGACCCAGTGTGGAAGCGGGGAGACATGAGGTGGA 327

Db 427 AAAACTCCTGGAAGGGAGGCCGTGTGAGGGCGGTCTGACCCAGTGTGGAAGCGGGGAGACATGAGGTGGA 486

Qy 328 TGGGCTCAATATATAACATTTGCGGTGAACCTTGATATTCCTAGATGCCAAAGGAAGATG 387

Db 487 TGGGCTCAATATATAACATTTGCGGTGAACCTTGATATTCCTAGATGCCAAAGGAAGATG 546

Qy 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTGTATCTGCTGATC 447

Db 547 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGGTGTATCTGCTGATC 606

Qy 448 CATATGTTTACAACTGGACAGCATGGTTCAGAGGACAGTACCGGGGAAAAATGGCACCGGCC 507

Db	607	CGTATGTTTACAACTGGACAGCATGGTCAGAGGACAGTGACGGGGAAAAATGGCACCGGCC	666
QY	508	AAAGCCATCATAAACGTCTTCCCTGATGGGAAACCTTTTCCCTCACCAACCCCGGATGGAGAA	567
Db	667	AAAGCCATCATAAACGTCTTCCCTGATGGGAAACCTTTTCCCTCACCAACCCCGGATGGAGAA	726
QY	568	GATGGAATTTTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT	627
Db	727	GATGGAATTTTCATCTACGTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT	786
QY	628	GTTCAGTGAGAGTTTCTGTGAAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG	687
Db	787	GTTCAGTGAGAGTTTCTGTGAAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG	846
QY	688	TGACTGTCTACAGAAAGACATGGACGGGCATATGTTCCCATCGCAAGTGAAGATGTGT	747
Db	847	TGACTGTCTACAGAAAGACATGGACGGGCATATGTTCCCATCGCAAGTGAAGATGTGT	906
QY	748	ACGTGGTAAACAGATCAGATTCCCTGTGTTTGTGACTATGTTCCAGAAACGATCGAAATT	807
Db	907	ACGTGGTAAACAGATCAGATTCCCTGTGTTTGTGACTATGTTCCAGAAACGATCGAAATT	966
QY	808	CATCCGACGAAACCTTCC-CAAAGATCTCCCCATATATGTTTGTATGTCTCTGATTCATGATC	866
Db	967	CATCCGACGAAACCTTCCCTCAAAGATCTCCCCATATATGTTTGTATGTCTCTGATTCATGATC	1026
QY	867	CTAGCCACTTCCCTCAATTATTCTACCAATTAACTACAAGTGGAGCTTCGGGGATAATACTG	926
Db	1027	CTAGCCACTTCCCTCAATTATTCTACCAATTAACTACAAGTGGAGCTTCGGGGATAATACTG	1086
QY	927	GCCTGTTTGTITCCACCAATCATACTGTGAATCACACGTATGTGCTCAATGGAACCTTCA	986
Db	1087	GCCTGTTTGTITCCACCAATCATACTGTGAATCACACGTATGTGCTCAATGGAACCTTCA	1146
QY	987	GCCTTAACCTCACTGTGAAAGCTGCAGCACCAAGACCTTGTCCGCCACCCGCCACCCAC	1046
Db	1147	GCCTTAACCTCACTGTGAAAGCTGCAGCACCAAGACCTTGTCCGCCACCCGCCACCCAC	1206
QY	1047	CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGCTGGERGACAACCCCTCGGAGCTGA	1106
Db	1207	CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGCTGGERGACAACCCCTCGGAGCTGA	1266
QY	1107	GTAGGATTCCTGATGAAACCTGCCAGATTAAACAGATATGSCCACTTTCAAGCCACCATCA	1166
Db	1267	GTAGGATTCCTGATGAAACCTGCCAGATTAAACAGATATGSCCACTTTCAAGCCACCATCA	1326
QY	1167	CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCTGATGCCGG	1226
Db	1327	CAATTGTAGAGGGAATCTTAGAGGTTAAACATCATCCAGATGACAGACGTCTGATGCCGG	1386
QY	1227	TGCCATGGCCTGAAAGCTCCCTAATAGACTTTTGTCTGACCTGCCAAGGGAGCATTTCCCA	1286
Db	1387	TGCCATGGCCTGAAAGCTCCCTAATAGACTTTTGTCTGACCTGCCAAGGGAGCATTTCCCA	1446
QY	1287	CGGAGGTCTGTACCATCAATTTCTGACCCCACTCGGAGATCACCCAGAACACAGCTGCA	1346
Db	1447	CGGAGGTCTGTACCATCAATTTCTGACCCCACTCGGAGATCACCCAGAACACAGCTGCA	1506
QY	1347	GCCCTGTGGATGTGGATGAGATGTGTCTGTGACTGTGAGACGAAACCTTCAATGGGTCTG	1406
Db	1507	GCCCTGTGGATGTGGATGAGATGTGTCTGTGACTGTGAGACGAAACCTTCAATGGGTCTG	1566
QY	1407	GGACGTACTGTGTGAACCTCACCCCTGGGGATGACAAAGCCTGGCTCTCACGAGCACCC	1466
Db	1567	GGACGTACTGTGTGAACCTCACCCCTGGGGATGACAAAGCCTGGCTCTCACGAGCACCC	1626
QY	1467	TGATTTCTGTTCCCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCTGA	1526
Db	1627	TGATTTCTGTTCCCTGACAGAGACCCAGCCTCGCCTTTAAGGATGGCAAAACAGTGCCTGA	1686
QY	1527	TCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAC	1586
Db	1687	TCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAC	1746

QY	1587	ACAAGGAATACAAACCCCAATAGAAAAATAGTCTCTGGGAATGTGGTCAGAAAGCAAGCCCTGA	1646
Db	1747	ACAAGGAATACAAACCCCAATAGAAAAATAGTCTCTGGGAATGTGGTCAGAAAGCAAGCCCTGA	1806
QY	1647	GTGTCTTTCTCAACCGTGC AAAAGCCGTGTTCTTCCCGGGAACCAAGGAAAGGATCCGC	1706
Db	1807	GTGTCTTTCTCAACCGTGC AAAAGCCGTGTTCTTCCCGGGAACCAAGGAAAGGATCCGC	1866
QY	1707	TACTCAAAAAACCAAGAAATTTTAAAGGAGTTTCTTAAATTTTCGACCTTGTCTGAAGCTCA	1766
Db	1867	TACTCAAAAAACCAAGAAATTTTAAAGGAGTTTCTTAAATTTTCGACCTTGTCTGAAGCTCA	1926
QY	1767	CTTTTCAGTGCCATTGATGTGAGATGTCTGGAGTGGCTATTAACCTTTTTCCTAAAG	1826
Db	1927	CTTTTCAGTGCCATTGATGTGAGATGTCTGGAGTGGCTATTAACCTTTTTCCTAAAG	1986
QY	1827	ATTATTGTTTAAATAGATAATTGTGTTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA	1886
Db	1987	ATTATTGTTTAAATAGATAATTGTGTTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA	2046
QY	1887	TTTTAGAGATGGGAGAGGGATTATACTGCAGGCAGCTTCAGCCATGTTGTGAAACTGAT	1946
Db	2047	TTTTAGAGATGGGAGAGGGATTATACTGCAGGCAGCTTCAGCCATGTTGTGAAACTGAT	2106
QY	1947	AAAAAGCAACTTAGCAAGGCTTCTTTTCATTATTTTATGTTTTCACTTATAAAGCTTTAG	2006
Db	2107	AAAAAGCAACTTAGCAAGGCTTCTTTTCATTATTTTATGTTTTCACTTATAAAGCTTTAG	2166
QY	2007	GTAACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGAGAGAAAGCTACTATTGATTA	2066
Db	2167	GTAACTAGTAGGATAGAAACACTGTGTCCCGAGAGTAAGAGAGAAAGCTACTATTGATTA	2226
QY	2067	GAGCCTAACCCAGGTTAACTGCAAGAAAGAGGCGGGATACTTTCAGCTTTCCTATGTAAGT	2126
Db	2227	GAGCCTAACCCAGGTTAACTGCAAGAAAGAGGCGGGATACTTTCAGCTTTCCTATGTAAGT	2286
QY	2127	TATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCA	2186
Db	2287	TATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCA	2346
QY	2187	CTTCAATACACACTCATGAACCTCCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGA	2246
Db	2347	CTTCAATACACACTCATGAACCTCCTGATGGAACAATAACAGGCCCAAGCCTGTGGTATGA	2406
QY	2247	TGTGCACACTTGTCTAGACTCAGAAAAATACTACTCTCATAAAATGGGTGGGAGTATTTTG	2306
Db	2407	TGTGCACACTTGTCTAGACTCAGAAAAATACTACTCTCATAAAATGGGTGGGAGTATTTTG	2466
QY	2307	GTGACAAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTATTTATTTCCA	2366
Db	2467	GTGACAAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTATTTATTTCCA	2526
QY	2367	TGGACATTTAGTTAGTCTTTTATATACCCAGGCATGATGCTGAGTGACACTTGTGTGA	2426
Db	2527	TGGACATTTAGTTAGTCTTTTATATACCCAGGCATGATGCTGAGTGACACTTGTGTGA	2586
QY	2427	TATTTCCAAAATTTTGTATAGTCGCTGCACATATTTGAAATCAAAATATTAAAGACTTTCC	2486
Db	2587	TATTTCCAAAATTTTGTACAGTCGCTGCACATATTTGAAATC-ATATATTAAAGACTTTCC	2645
QY	2487	AAAAATTTGGTCCCTGGTTTTTTCATGGCAACTGATCAGTAAGGATTTCCCTCTGTGTTG	2546
Db	2646	AAAGATGAGGTCCTCGTTTTTTCATGGCAACTGATCAGTAAGGATTTTCACCTCTGTGTTG	2705
QY	2547	GAACTAAAAACCATTTACTATATGTTTAGACAAGACATTTTTTTTTTTTCTCTCTGAAAAA	2606
Db	2706	TAACTAAAAACCATTTACTATATGTTTAGACAAGACATTTCTTTTCTCTCTCTCTGAAAAA	2765
QY	2607	-AAAATGAGGGAAGAGACAAAAA AAAAAAAAAAAAAAAAAAAAAA 2657	
Db	2766	TAAAGTGTGGGAAGAGACAAAAA AAAAAAAAAAAAAAACC AAAAGAAAGAA 2817	


```
QY 1767 CTTTTCAGTCCATTGATGTGAGATGTGCTGGAGTGGCTATTAACTTTTTCCTAAAG 1826
Db 1800 CTTTTCAGTCCATTGATGTGAGATGTGCTGGAGTGGCTATTAACTTTTTCCTAAAG 1859
QY 1827 ATTATTTGTTAAATAGATATTGTGTTGGGGAAGTGAATTTTATAGTTAAATGTCA 1886
Db 1860 ATTATTTGTTAAATAGATATTGTGTTGGGGAAGTGAATTTTATAGTTAAATGTCA 1919
QY 1887 TTTTAGAGATGGGAGAGGAGATTATATCTGACAGCAGCTTCAGCCATGTGGAAGTGAAT 1946
Db 1920 TTTTAGAGATGGGAGAGGAGATTATATCTGACAGCAGCTTCAGCCATGTGGAAGTGAAT 1979
QY 1947 AAAAGCACTTAGCAAGGCTTCTTTTCATTATTTTATATGTTTCACTTAATAAGTCTTAG 2006
Db 1980 AAAAGCACTTAGCAAGGCTTCTTTTCATTATTTTATATGTTTCACTTAATAAGTCTTAG 2039
QY 2007 GTAACTAGTAGGATGAAACACTGTGTCCCGAGAGTAAGAGAGAGAGTAACTATTGATTA 2066
Db 2040 GTAACTAGTAGGATGAAACACTGTGTCCCGAGAGTAAGAGAGAGAGTAACTATTGATTA 2099
QY 2067 GAGCCTAACCCAGGTTAACTGCAAGAGAGGCGGATACCTTTCAGCTTTCATGTAAGTGA 2126
Db 2100 GAGCCTAACCCAGGTTAACTGCAAGAGAGGCGGATACCTTTCAGCTTTCATGTAAGTGA 2159
QY 2127 TATGCAATAAGCCAAATGATGTCAGTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCA 2186
Db 2160 TATGCAATAAGCCAAATGATGTCAGTTTCTAAGATCATGTTCCAAGCTAACTGAATCCCA 2219
QY 2187 CTTCAATACACACTCATGAACCTCCTGATGGAACAATAACAGGCCCAAGCCTGTGATGTA 2246
Db 2220 CTTCAATACACACTCATGAACCTCCTGATGGAACAATAACAGGCCCAAGCCTGTGATGTA 2279
QY 2247 TGTGACACACTTGTGTAAGTCTGAGAAATAATACTACTCTCATAAATGGGTGGAGTATTGTG 2306
Db 2280 TGTGACACACTTGTGTAAGTCTGAGAAATAATACTACTCTCATAAATGGGTGGAGTATTGTG 2339
QY 2307 GTGACAACCTACTTGTGCTGGCTGAGTGAAGGAATGATATTCAATATTCATTATTCCA 2366
Db 2340 GTGACAACCTACTTGTGCTGGCTGAGTGAAGGAATGATATTCAATATTCATTATTCCA 2399
QY 2367 TGGACATTTAGTTAGTCTTTTATATACAGGAGATGATGCTGAGTACACTCTTGTTGA 2426
Db 2400 TGGACATTTAGTTAGTCTTTTATATACAGGAGATGATGCTGAGTACACTCTTGTTGA 2459
QY 2427 TATTTCCAAATTTTGTATAGTGTGCTGACATATTTGAATCAAAATATTAAGACTTCC 2486
Db 2460 TATTTCCAAATTTTGTATAGTGTGCTGACATATTTGAATC-ATAATATTAAGACTTCC 2518
QY 2487 AAAAATTTGGTCCCTGGTTTTCATGGCAACTGATCAGTAAGGATTTCCCTGTGTTG 2546
Db 2519 AAAAGATGAGTCCCTGGTTTTCATGGCAACTGATCAGTAAGGATTTCACTCTGTTTG 2578
QY 2547 GAACTAAAACCATTTACTATATGTTAGCAAGACATTTTATTTTCTCTCTGAAAAA 2606
Db 2579 TAACTAAAACCATTTACTATATGTTAGCAATTCATTTTCTCTCTCTGAAAAA 2638
QY 2607 -AAAAATGAGGAGAGAGACAAAAA 2636
Db 2639 TAAAGTGTGGGAGAGACAAAAA 2669

RESULT 5
US-10-342-887-779
; Sequence 779, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yundong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
```

```
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 779
; LENGTH: 2669
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-779

Query Match          95.7%; Score 2547.8; DB 13; Length 2669;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 28 AACCTGTGTCCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTTACTATTTCCTGGAT 87
Db 60 ACCTTGATGTCCTGCGTCCGTGAGAAATTCAGCATGGAATGTCTTACTATTTCCTGGAT 119
QY 88 TTCTGCTCTGCTGCAAGATTGCCATTGATGCCGCCAAACGATTTCATGATGTGCTGG 147
Db 120 TTCTGCTCTGCTGCAAGATTGCCATTGATGCCGCCAAACGATTTCATGATGTGCTGG 179
QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGAGCAACAATCAATTAATGCTGCTTCTG 207
Db 180 GCAATGAAAGACCTTCTGCTTACATGAGGAGCAACAATCAATTAATGCTGCTTCTG 239
QY 208 ATGAAATGACTGGAATGAAAACTTACCAGTGTGGAAGCGGGAGACATGAGGTGA 267
Db 240 ATGAAATGACTGGAATGAAAACTTACCAGTGTGGAAGCGGGAGACATGAGGTGA 299
QY 268 AAAACTCTGGAAGGAGGCGGTGTGAGAGCGGCTCTGACCAAGTGACTCACAGCCCTCG 327
Db 300 AAAACTCTGGAAGGAGGCGGTGTGAGAGCGGCTCTGACCAAGTGACTCACAGCCCTCG 359
QY 328 TGGGCTCAAAATATAACATTTGCGGTGAACCTGATATTCCCTAGATGCCAAAAAGGAATG 387
Db 360 TGGGCTCAAAATATAACATTTGCGGTGAACCTGATATTCCCTAGATGCCAAAAAGGAATG 419
QY 388 CCAATGGCAACATAGTCTATGAGAGAACTGCAGAAATGAGCGCTGTTATCTGCTGATC 447
Db 420 CCAATGGCAACATAGTCTATGAGAGAACTGCAGAAATGAGCGCTGTTATCTGCTGATC 479
QY 448 CATATGTTTACAACGTGACAGACATGCTCAGAGACAGTGAACGGGAAAAATGCAACCGGCC 507
Db 480 CATATGTTTACAACGTGACAGACATGCTCAGAGACAGTGAACGGGAAAAATGCAACCGGCC 539
QY 508 AAAGCCATCATAAACGCTTCCCTGATGGGAAAAACCTTTTCTCTCACACACCCCGATGAGAA 567
Db 540 AAAGCCATCATAAACGCTTCCCTGATGGGAAAAACCTTTTCTCTCACACACCCCGATGAGAA 599
QY 568 GATGCAATTTCAATCTACGTCTTCCACACACTTGGTCAAGTATTTCCAGAAATTTGGGACGAT 627
Db 600 GATGCAATTTCAATCTACGTCTTCCACACACTTGGTCAAGTATTTCCAGAAATTTGGGACGAT 659
QY 628 GTTCAGTAGAGTTTCTGTGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAG 687
Db 660 GTTCAGTAGAGTTTCTGTGAACACAGCCAAATGTGACACTTGGGCTCAACTCATGGAAG 719
QY 688 TGACTGTCTACAGAGACATGAGCGGCATATGTTCCCATCGCAAGTGAAGATGTGT 747
Db 720 TGACTGTCTACAGAGACATGAGCGGCATATGTTCCCATCGCAAGTGAAGATGTGT 779
QY 748 ACGTGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAAGCATGGAATTT 807
Db 780 ACGTGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAAGCATGGAATTT 839
```

```
QY      808 CATCCGACGAAACCTTCC-CAAGATCTCCCATTTATGTTTGATGTCTGATTCATGATC 866
      |||
Db      840 CATCCGACGAAACCTTCCCAAGATCTCCCATTTATGTTTGATGTCTGATTCATGATC 899
QY      867 CTAGCCACTTCTCAATTTATTTCTACATTTAACAAGTGAGCTTCGGGATTAATACTG 926
      |||
Db      900 CTAGCCACTTCTCAATTTATTTCAACATTTAACAAGTGAGCTTCGGGATTAATACTG 959
QY      927 GCCTGTTGTTTCCACCAATCATCTGTGAATCACACGTAATGTGCTCAATGGAACCTTCA 986
      |||
Db      960 GCCTGTTGTTTCCACCAATCATCTGTGAATCACACGTAATGTGCTCAATGGAACCTTCA 1019
QY      987 GCCTTAACCTCACTGTGAAGCTGCAGACCAAGACCTTGTCCGCCAACCACCACAC 1046
      |||
Db      1020 GCCTTAACCTCACTGTGAAGCTGCAGACCAAGACCTTGTCCGCCAACCACCACAC 1079
QY      1047 CCAGACCTTCAAAACCCACCCCTTTCTTTAGGACCTGCTGTGACAACCCCTGAGCTGA 1106
      |||
Db      1080 CCAGACCTTCAAAACCCACCCCTTTCTTTAGGACCTGCTGTGACAACCCCTGAGCTGA 1139
QY      1107 GTAGGATTCCTGATGAAAACTGCCAGATTTAACAGATATGCCACTTCAAGCCACCATCA 1166
      |||
Db      1140 GTAGGATTCCTGATGAAAACTGCCAGATTTAACAGATATGCCACTTCAAGCCACCATCA 1199
QY      1167 CAATTTAGAGGGAATCTTAGAGGTTAACATCATCCAGATGACAGACGTCCTGATGCCG 1226
      |||
Db      1200 CAATTTAGAGGGAATCTTAGAGGTTAACATCATCCAGATGACAGACGTCCTGATGCCG 1259
QY      1227 TGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGCCAAGGAGCATTTCCA 1286
      |||
Db      1260 TGCCATGGCCTGAAAGCTCCCTAATAGACTTTGTGCTGACCTGCCAAGGAGCATTTCCA 1319
QY      1287 CGGAGGTCTGTACCATATTTCTGACCCCACTCGGAGATCACCAGAACACAGTCTGCA 1346
      |||
Db      1320 CGGAGGTCTGTACCATATTTCTGACCCCACTCGGAGATCACCAGAACACAGTCTGCA 1379
QY      1347 GCCCTGTGATGTGATGAGATGTGTCTGCTGACTGTGAGAGAACTTCAATGGGTCTG 1406
      |||
Db      1380 GCCCTGTGATGTGATGAGATGTGTCTGCTGACTGTGAGAGAACTTCAATGGGTCTG 1439
QY      1407 GGAAGTACTGTGTGAACCTCACTGGGGGATGACACAAGCTGCTCTCAGAGCACCC 1466
      |||
Db      1440 GGAAGTACTGTGTGAACCTCACTGGGGGATGACACAAGCTGCTCTCAGAGCACCC 1499
QY      1467 TGAATTTCTGTTCTGACAGAGACCACTCGCTTTAAGGATGGCAAAACAGTGCCTGCA 1526
      |||
Db      1500 TGAATTTCTGTTCTGACAGAGACCACTCGCTTTAAGGATGGCAAAACAGTGCCTGCA 1559
QY      1527 TCTCCGTTGGCTGCTTGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAC 1586
      |||
Db      1560 TCTCCGTTGGCTGCTTGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAC 1619
QY      1587 ACAAGGAATACAACCAATAGAAATAAGTCTCTGGGAATGTGTCAAGCAAGGCTCTGA 1646
      |||
Db      1620 ACAAGGAATACAACCAATAGAAATAAGTCTCTGGGAATGTGTCAAGCAAGGCTCTGA 1679
QY      1647 GTGTCTTTCTCAACCGTGCAAAAGCCGTCTTCTCCCGGGAACCAAGAAAAGATCCGC 1706
      |||
Db      1680 GTGTCTTTCTCAACCGTGCAAAAGCCGTCTTCTCCCGGGAACCAAGAAAAGATCCGC 1739
QY      1707 TACTCAAAAAACAAGATTAAAGAGTTTCTTAAATTTCGACCTGTCTTCTGAAGCTCA 1766
      |||
Db      1740 TACTCAAAAAACAAGATTAAAGAGTTTCTTAAATTTCGACCTGTCTTCTGAAGCTCA 1799
QY      1767 CTTTCAAGTGCATTTGATGTGAGATGTGCTGAGTGGCTATTAACCTTTTTCCTTAAG 1826
      |||
Db      1800 CTTTCAAGTGCATTTGATGTGAGATGTGCTGAGTGGCTATTAACCTTTTTCCTTAAG 1859
QY      1827 ATTAATGTTAAATAGATATTTGTGTTGGGGAAGTGAATTTTATAGGTTAATGTCA 1886
      |||
Db      1860 ATTAATGTTAAATAGATATTTGTGTTGGGGAAGTGAATTTTATAGGTTAATGTCA 1919
```

```
QY      1887 TTTTAGAGATGGGAGAGGGAATTATCTGACAGGCACTTCAGCCATGTTGTGAAACTGAT 1946
      |||
Db      1920 TTTTAGAGATGGGAGAGGGAATTATCTGACAGGCACTTCAGCCATGTTGTGAAACTGAT 1979
QY      1947 AAAAGCACTTAGCAAGCTTCTTTTCATTTATTTTATAGTTTCACTTATAAGTCTTAG 2006
      |||
Db      1980 AAAAGCACTTAGCAAGCTTCTTTTCATTTATTTTATAGTTTCACTTATAAGTCTTAG 2039
QY      2007 GTAAGTAGTAGTAGTAAGAACACTGTGTCCCGAGATTAAGAGAGAGAGTACTATTTGATTA 2066
      |||
Db      2040 GTAAGTAGTAGTAGTAAGAACACTGTGTCCCGAGATTAAGAGAGAGTACTATTTGATTA 2099
QY      2067 GAGCCTAACCCAGGTTAAGTGAAGAGAGAGGCGGATATTTCAGCTTTCCATGTAACTG 2126
      |||
Db      2100 GAGCCTAACCCAGGTTAAGTGAAGAGAGAGGCGGATATTTCAGCTTTCCATGTAACTG 2159
QY      2127 TATGATTAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAAAGCTAACTGAATCCA 2186
      |||
Db      2160 TATGATTAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAAAGCTAACTGAATCCA 2219
QY      2187 CTTCAATACACACTCATGAATCTCTGATGGAACAAATTAACAGGCCCAAGCCTGTGTATGA 2246
      |||
Db      2220 CTTCAATACACACTCATGAATCTCTGATGGAACAAATTAACAGGCCCAAGCCTGTGTATGA 2279
QY      2247 TGTGCACACTTGCTAGACTGCAAAAAAATACTACTCTCATATAATGGGTGGAGTATTTG 2306
      |||
Db      2280 TGTGCACACTTGCTAGACTGCAAAAAAATACTACTCTCATATAATGGGTGGAGTATTTG 2339
QY      2307 GTGACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATTTCA 2366
      |||
Db      2340 GTGACAACCTACTTTGCTTGGCTGAGTGAAGGAATGATATTCATATATTCATTTATTTCA 2399
QY      2367 TGGACATTTAGTTAGTGTCTTTTATATATACAGGCAATGCTGAGTGACACTCTGTGTGA 2426
      |||
Db      2400 TGGACATTTAGTTAGTGTCTTTTATATATACAGGCAATGCTGAGTGACACTCTGTGTGA 2459
QY      2427 TATTTCCAAATTTTGTATAGTGCCTGACATATTTGAAATCAAAATATTAAGACTTTCC 2486
      |||
Db      2460 TATTTCCAAATTTTGTATAGTGCCTGACATATTTGAAATC-ATATATTAAGACTTTCC 2518
QY      2487 AAAAATTTGTCCTGTTTTTTCATGGCAACTGATCAATAAGATTTTCCCTCTGTTTG 2546
      |||
Db      2519 AAAGATGAGTCCCTGTTTTTCATGGCAACTGATCAATAAGATTTCACTCTGTTTG 2578
QY      2547 GAACTAAAAACCATTTACTATATGTTAGACAAGACATTTTTTTTTCTCTCTGGAATAA 2606
      |||
Db      2579 TAACTAAAAACCATTTACTATATGTTAGACATGACATTTCTTTCTCTCTCTGAAAAA 2638
QY      2607 -AAAAAGAAGGAGACAAAAA 2636
      |||
Db      2639 TAAAGTGTGGGAAGACAAAAA 2669
```

```
RESULT 6
US-10-172-118-779
: Sequence 779, Application US/10172118
: Publication No. US20030224374A1
: GENERAL INFORMATION:
: APPLICANT: Dai, Hongyue
: APPLICANT: He, Yudong
: APPLICANT: Linsley, Peter
: APPLICANT: Mao, Mao
: APPLICANT: Roberts, Chris
: APPLICANT: Van 't Veer, Laura
: APPLICANT: Van de Vijver, Marc
: APPLICANT: Bernards, Rene
: TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
: FILE REFERENCE: 9301-175-999
: CURRENT APPLICATION NUMBER: US/10/172,118
: CURRENT FILING DATE: 2002-06-14
: PRIOR APPLICATION NUMBER: 60/380,770
: PRIOR FILING DATE: 2002-05-14
: NUMBER OF SEQ ID NOS: 2699
```


; SEQ ID NO 779
; LENGTH: 2669
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATON INFORMATION:
; DATABASE ACCESSION NUMBER: NM 002510
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-779

Query Match 95.7%; Score 2547.8; DB 13; Length 2669;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 28 AACCTGGTGGCTGGCTCCGTGAGAATTGAGATGGAATGTCCTACTAATTTCCGGGAT 87
Db 60 ACCTGAGTGGCTGGCTCCGTGAGAATTGAGATGGAATGTCCTACTAATTTCCGGGAT 119
QY 88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTCATGATGTGCTGG 147
Db 120 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTCATGATGTGCTGG 179
QY 148 GCAATGAAGAACCCTTCTGCTTACATGAGGAGACACAATCAATTAATGCGCTGCTTCTG 207
Db 180 GCAATGAAGAACCCTTCTGCTTACATGAGGAGACACAATCAATTAATGCGCTGCTTCTG 239
QY 208 ATGAATAATGACTGGAATGAAAACTCTAACCCAGTGTGGAAGCGGGAGACATGAGGTGA 267
Db 240 ATGAATAATGACTGGAATGAAAACTCTAACCCAGTGTGGAAGCGGGAGACATGAGGTGA 299
QY 268 AAAACTCCTGGAAGGAGGAGCGCGGTGTGGAAGCGCGTCCGTGACCAATGACTCACCAGCCCTCG 327
Db 300 AAAACTCCTGGAAGGAGGAGCGCGGTGTGGAAGCGCGTCCGTGACCAATGACTCACCAGCCCTCG 359
QY 328 TGGGCTCAATATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG 387
Db 360 TGGGCTCAATATATAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG 419
QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAATGAGCGCTGTTATCTGCTGATC 447
Db 420 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAATGAGCGCTGTTATCTGCTGATC 479
QY 448 CATATGTTTACAACCTGAGACAGCATGCTCAGAGACAGTGAACGGGAAATGGCACCGGCC 507
Db 480 CATATGTTTACAACCTGAGACAGCATGCTCAGAGACAGTGAACGGGAAATGGCACCGGCC 539
QY 508 AAAGCCATCATACGTCCTCCCTGATGGGAAACCTTTTCTCTCAACACCCCGGATGAGAA 567
Db 540 AAAGCCATCATACGTCCTCCCTGATGGGAAACCTTTTCTCTCAACACCCCGGATGAGAA 599
QY 568 GATGGAATTTTCATCTACGTCCTCCACACACACTGTGTCAGTATTTCCAGAAATTTGGGACGAT 627
Db 600 GATGGAATTTTCATCTACGTCCTCCACACACACTGTGTCAGTATTTCCAGAAATTTGGGACGAT 659
QY 628 GTTCAGTGAGAGTTTCTGTGAACAACAGCCAATGTGACACTTGGGCCTCAACTCATGGAAG 687
Db 660 GTTCAGTGAGAGTTTCTGTGAACAACAGCCAATGTGACACTTGGGCCTCAACTCATGGAAG 719
QY 688 TGAAGTGTACAGAGACATGAGACGGGACATATGTTCCCATCGCACAAAGTGAAGATGTGT 747
Db 720 TGAAGTGTACAGAGACATGAGACGGGACATATGTTCCCATCGCACAAAGTGAAGATGTGT 779
QY 748 ACGTGTACAGAGACATGAGATTCCTGTGTGTTGTGACTATGTTCCAGAAAGCATCGAAATT 807
Db 780 ACGTGTACAGAGACATGAGATTCCTGTGTGTTGTGACTATGTTCCAGAAAGCATCGAAATT 839
QY 808 CATCCGAGAAACCTTCC-CAAAGATCTCCCATTAATGTTGATGCTGATTCATGATC 866
Db 840 CATCCGAGAAACCTTCC-CAAAGATCTCCCATTAATGTTGATGCTGATTCATGATC 899
QY 867 CTAGCCACTTCTCAATTAATTTCTACCATTAACAAAGTGAGCTTCGGGATAATACTG 926
Db 900 CTAGCCACTTCTCAATTAATTTCTACCATTAACAAAGTGAGCTTCGGGATAATACTG 959

QY 927 GCCTGTTGTTTCCACCAATCATACTGTGAATCACAACGATGTGCTCAATGGAACCTTCA 986
Db 960 GCCTGTTGTTTCCACCAATCATACTGTGAATCACAACGATGTGCTCAATGGAACCTTCA 1019
QY 987 GCCTTAACCTCACTGTGAAGCTGCAGACACAGGACCTTGTCCGCCACCGCACCAACAC 1046
Db 1020 GCCTTAACCTCACTGTGAAGCTGCAGACACAGGACCTTGTCCGCCACCGCACCAACAC 1079
QY 1047 CCAGACCTTCAAAAACCCACCCCTTCTTTAGGACCTGCTGTGACACAACCCCTGAGCTGA 1106
Db 1080 CCAGACCTTCAAAAACCCACCCCTTCTTTAGGACCTGCTGTGACACAACCCCTGAGCTGA 1139
QY 1107 GTAGATTCCTGATGAAAACTGCCAGATTAACAGATATGGCCACTTCAACCAACATCA 1166
Db 1140 GTAGATTCCTGATGAAAACTGCCAGATTAACAGATATGGCCACTTCAACCAACATCA 1199
QY 1167 CAATTGTAGAGGGAATCTTAGAGGTTAACATCATCCAGATGACAGAGTCTGATGCCGG 1226
Db 1200 CAATTGTAGAGGGAATCTTAGAGGTTAACATCATCCAGATGACAGAGTCTGATGCCGG 1259
QY 1227 TGCCATGGCCCTGAAAGCTCCCTAATAGACTTTGTGCTGAGACCTGCCAAGGAGCATTCOCA 1286
Db 1260 TGCCATGGCCCTGAAAGCTCCCTAATAGACTTTGTGCTGAGACCTGCCAAGGAGCATTCOCA 1319
QY 1287 CGAGGTCTGTACCATCATTTCTGACCCCACTCGAGATGACCCAGAACAACAGTCTGCA 1346
Db 1320 CGAGGTCTGTACCATCATTTCTGACCCCACTCGAGATGACCCAGAACAACAGTCTGCA 1379
QY 1347 GCCCTGTGATGTGATGAGATGATGCTGCTGACTGTGAGACGAACCTTCAATGGGTCTG 1406
Db 1380 GCCCTGTGATGTGATGAGATGATGCTGCTGACTGTGAGACGAACCTTCAATGGGTCTG 1439
QY 1407 GGAAGTACTGTGTGAACCTCAACCTGGGGGATGACACAAGCTGCTCTCAGAGACACCC 1466
Db 1440 GGAAGTACTGTGTGAACCTCAACCTGGGGGATGACACAAGCTGCTCTCAGAGACACCC 1499
QY 1467 TGAATTTCTGTTCTGACAGAGACCCAGCCTGCGCTTTAAGATGGCAACAGTGCCTGA 1526
Db 1500 TGAATTTCTGTTCTGACAGAGACCCAGCCTGCGCTTTAAGATGGCAACAGTGCCTGA 1559
QY 1527 TCTCCGTTGGCTGCTGGCCATATTTGTCACTGTGATCCCTTGTGTTACAAAAAAC 1586
Db 1560 TCTCCGTTGGCTGCTGGCCATATTTGTCACTGTGATCCCTTGTGTTACAAAAAAC 1619
QY 1587 ACAAGGAATACAAACCCATAAGAAATAGTCTGGGAATGTGTCAGAGCAAAAGGCTGA 1646
Db 1620 ACAAGGAATACAAACCCATAAGAAATAGTCTGGGAATGTGTCAGAGCAAAAGGCTGA 1679
QY 1647 GTGTCTTCTCAACCGTGCAAAAAGCCGTGTTCTTCCGGGAAACCAAGAAAGATCCGC 1706
Db 1680 GTGTCTTCTCAACCGTGCAAAAAGCCGTGTTCTTCCGGGAAACCAAGAAAGATCCGC 1739
QY 1707 TACTCAAAAACCAAGAAATTTAAAGAGTTCTTAAATTTGCACTTGTTCGAAGCTCA 1766
Db 1740 TACTCAAAAACCAAGAAATTTAAAGAGTTCTTAAATTTGCACTTGTTCGAAGCTCA 1799
QY 1767 CTTTTCAGTCCATGATGTGAGATGTGCTGAGTGGCTATTAAACCTTTTCTCTAAAG 1826
Db 1800 CTTTTCAGTCCATGATGTGAGATGTGCTGAGTGGCTATTAAACCTTTTCTCTAAAG 1859
QY 1827 ATTATTGTAAATAGATATTTGTGTTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA 1886
Db 1860 ATTATTGTAAATAGATATTTGTGTTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA 1919
QY 1887 TTTTAGAGATGGGAGAGAGGATTTATACCTGACGCACTTCAGCCATGTTGTGAACCTGAT 1946
Db 1920 TTTTAGAGATGGGAGAGAGGATTTATACCTGACGCACTTCAGCCATGTTGTGAACCTGAT 1979
QY 1947 AAAAGCACTTAGCAAGGCTCTTTTCAATATTTTATATGTTTCACTTAATAAGTCTTAG 2006
Db 1980 AAAAGCACTTAGCAAGGCTCTTTTCAATATTTTATATGTTTCACTTAATAAGTCTTAG 2039
QY 2007 GTAACTAGTAGATAGAAAACACTGTGTCCGAGAGTAAGAGAGAGAAAGCTAATTGATTA 2066

Db 780 ACGTGGTAACAGATCAGATTCCGTGTGTAAGTATGTTCCAGAGAAACGATCGAAATT 839
QY 808 CATCCGACGAAACCTTCC-CAAGAATCTCCCATTAATGTTGATGTCTGATTCATGATC 866
Db 840 CATCCGACGAAACCTTCTCAAGATCTCCCATTAATGTTGATGTCTGATTCATGATC 899
QY 867 CTAGCCACTTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 926
Db 900 CTAGCCACTTCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 959
QY 927 GCCTGTTGTTCCAAATCATATCTGTAATCAACGTATGTGCTCAATGGAACCTTCA 986
Db 960 GCCTGTTGTTCCAAATCATATCTGTAATCAACGTATGTGCTCAATGGAACCTTCA 1019
QY 987 GCCTTAACCTCACTGTGAAGCTGCAGCACCAGAACCTTGTCGCCACCGCCACACCCAC 1046
Db 1020 GCCTTAACCTCACTGTGAAGCTGCAGCACCAGAACCTTGTCGCCACCGCCACACCCAC 1079
QY 1047 CCAGACCTTCAAAACCCACCCCTTCTTGAAGCCTGCTGTGACAAACCCCTGGAGCTGA 1106
Db 1080 CCAGACCTTCAAAACCCACCCCTTCTTGAAGCCTGCTGTGACAAACCCCTGGAGCTGA 1139
QY 1107 GTAGGATTCCTGATGAAAACTGCCAGATTAACAGATATGGCCACTTGAAGCCACCATCA 1166
Db 1140 GTAGGATTCCTGATGAAAACTGCCAGATTAACAGATATGGCCACTTGAAGCCACCATCA 1199
QY 1167 CAATTGTAGAGGAACTTTAGAGTTAACTCATCCAGATGACAGACGTCTGATGCCGG 1226
Db 1200 CAATTGTAGAGGAACTTTAGAGTTAACTCATCCAGATGACAGACGTCTGATGCCGG 1259
QY 1227 TGCCATGGCCTGAAAAAGCTCCCTAATAGACTTTGTGCGTGAACCTGCCAAGGAGCATTTCCA 1286
Db 1260 TGCCATGGCCTGAAAAAGCTCCCTAATAGACTTTGTGCGTGAACCTGCCAAGGAGCATTTCCA 1319
QY 1287 CGGAGGTCTGTACCATCATTTTCTGACCCCACTGCGAGATCACCCAAGAACACAGTCTGCA 1346
Db 1320 CGGAGGTCTGTACCATCATTTTCTGACCCCACTGCGAGATCACCCAAGAACACAGTCTGCA 1379
QY 1347 GCCCTGTGATGTGATGAGATGTGTCTGTGACTGTGAGACGAACCTTCAATGGGTGTG 1406
Db 1380 GCCCTGTGATGTGATGAGATGTGTCTGTGACTGTGAGACGAACCTTCAATGGGTGTG 1439
QY 1407 GGACGTACTGTGTGAACCTCAACCTGGGGATGACACAAGCCTGGCTCTCAAGACACCC 1466
Db 1440 GGACGTACTGTGTGAACCTCAACCTGGGGATGACACAAGCCTGGCTCTCAAGACACCC 1499
QY 1467 TGATTTCTGTCTGACAGAGACCAGCCCTGCTTTAAGATGGCAACAGTGCCTGA 1526
Db 1500 TGATTTCTGTCTGACAGAGACCAGCCCTGCTTTAAGATGGCAACAGTGCCTGA 1559
QY 1527 TCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAAC 1586
Db 1560 TCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTCTTGGTGTACAAAAAAC 1619
QY 1587 ACAAGGAATACAAACCAATAGAAAAATAGTCTGGGAATGTGTACAGAACAAAGGCCTGA 1646
Db 1620 ACAAGGAATACAAACCAATAGAAAAATAGTCTGGGAATGTGTACAGAACAAAGGCCTGA 1679
QY 1647 GTGTCTTTCTCAACCGTGCAAAAGCCGTGTCTTCCCGGAAACCAAGAAAAAGATCCGC 1706
Db 1680 GTGTCTTTCTCAACCGTGCAAAAGCCGTGTCTTCCCGGAAACCAAGAAAAAGATCCGC 1739
QY 1707 TACTCAAAAACCAAGATTTTAAAGAGTTTCTTAAATTTGACCTTGTCTGAAGCTCA 1766
Db 1740 TACTCAAAAACCAAGATTTTAAAGAGTTTCTTAAATTTGACCTTGTCTGAAGCTCA 1799
QY 1767 CTTTTCAGTGCATTTGATGTGAGATGTGTGAGATGGCTATTAACCTTTTTTCTCTAAG 1826
Db 1800 CTTTTCAGTGCATTTGATGTGAGATGTGTGAGATGGCTATTAACCTTTTTTCTCTAAG 1859
QY 1827 ATTATGTTAATAGATATTTGTGTTGGGGAAGTTGAATTTTTTATAGCTTAAATGTCA 1886

Db 1860 ATTATGTTAATAGATATTTGTGTTTGGGAAAGTTGAATTTTTTATAGTTAAATGTCA 1919
QY 1887 TTTTAGATGGGGAGAGGATTAATCTGACAGCAGCTTCAGCCATGTGTGAACATGAT 1946
Db 1920 TTTTAGATGGGGAGAGGATTAATCTGACAGCAGCTTCAGCCATGTGTGAACATGAT 1979
QY 1947 AAAAGCACTTAGCAAGGCTCTTTTCATTATTTTATGTCTACTTATPAAGTCTTAG 2006
Db 1980 AAAAGCACTTAGCAAGGCTCTTTTCATTATTTTATGTCTACTTATPAAGTCTTAG 2039
QY 2007 GTAACTAGAGATAGAAACACTGTGTCCCGAGAGTAAGAGAGAGACTAATTTGATTA 2066
Db 2040 GTAACTAGAGATAGAAACACTGTGTCCCGAGAGTAAGAGAGAGACTAATTTGATTA 2099
QY 2067 GAGCCTAACCCAGGTTAACTGCAGAGAGAGCGGATACTTTGAGCTTCCATGTAACGTG 2126
Db 2100 GAGCCTAACCCAGGTTAACTGCAGAGAGAGCGGATACTTTGAGCTTCCATGTAACGTG 2159
QY 2127 TATGCAATAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCAGCTAACGTAATCCCA 2186
Db 2160 TATGCAATAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCAGCTAACGTAATCCCA 2219
QY 2187 CTTCAATACACACTCATGAACTCTCTGATGGAACAATPAACAGGCCCAAGCCTGTGTATGA 2246
Db 2220 CTTCAATACACACTCATGAACTCTCTGATGGAACAATPAACAGGCCCAAGCCTGTGTATGA 2279
QY 2247 TGTGACACTGTCTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGAGTATTTTG 2306
Db 2280 TGTGACACTGTCTAGACTCAGAAAAAATACTACTCTCATAAATGGGTGGAGTATTTTG 2339
QY 2307 GTGACAACTACTTTGCTTGGCTGAGTGAAGGAATGATTCATATTTCAATTTATTTCCA 2366
Db 2340 GTGACAACTACTTTGCTTGGCTGAGTGAAGGAATGATTCATATTTCAATTTATTTCCA 2399
QY 2367 TGACACTTAGTTAGTGTCTTTTATATACAGGCATGATGCTGAGTACACTCTGTGTGA 2426
Db 2400 TGACACTTAGTTAGTGTCTTTTATATACAGGCATGATGCTGAGTACACTCTGTGTGA 2459
QY 2427 TATTTCCAAATTTTGTATAGTGGCTGCACATATTTGGAATCAAAATATTAAGACTTCC 2486
Db 2460 TATTTCCAAATTTTGTATAGTGGCTGCACATATTTGGAATC-ATATATTAAGACTTCC 2518
QY 2487 AAAAATTTGTCCTCGTTTTCATGCGCAACTTGATCAAGTAAGATTTCCCTGTGTTG 2546
Db 2519 AAAGATGAGTCCCTGTTTTCATGCGCAACTTGATCAAGTAAGATTTCACTCTGTTTG 2578
QY 2547 GAACTAAAAACATTTACTATATGTGTAGCAAGACATTTTTTTTTTCTCTGAAAAA 2606
Db 2579 TAACTAAAAACATCTACTATATGTGTAGACATGACATTTCTTTCTCTCTGAAAAA 2638
QY 2607 -AAAAAGGGAAGAGACAAAAA 2636
Db 2639 TAAAGTGTGGGAAGAGACAAAAA 2669

RESULT 8
US-10-450-826-92
; Sequence 92, Application US/10450826
; Publication No. US20040101818A1
; GENERAL INFORMATION:
; APPLICANT: Ji, Darren
; APPLICANT: Axelrod, Douglas W.
; APPLICANT: Cook, Jonathan S.
; APPLICANT: Jaiswal, Neelam
; APPLICANT: Eistein, Richard
; APPLICANT: Houghton, Adam
; APPLICANT: Mertz, Lawrence
; TITLE OF INVENTION: Gene Expression Profiles Associated with Osteoblast Differentiation
; FILE REFERENCE: 044921-5039-WO
; CURRENT APPLICATION NUMBER: US/10/450,826
; CURRENT FILING DATE: 2003-06-18
; PRIOR APPLICATION NUMBER: US 60/255,882
; PRIOR FILING DATE: 2000-12-18

; PRIOR APPLICATION NUMBER: US 60/285,691
; PRIOR FILING DATE: 2001-04-24
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
; LENGTH: 2669
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. X76534
US-10-450-826-92

Query Match 95.7%; Score 2547.8; DB 17; Length 2669;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2591; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 28 AACCTTGCTGCTGCGTCCGTGAGATTGACATGGAATGCTCTACTATTTCCTGGGAT 87
 60 ACCTTGAGTGCTGCGTCCGTGAGATTGACATGGAATGCTCTACTATTTCCTGGGAT 119
QY 88 TTCTGCTCCTGGCTGCAAGATTGCCAATTGATGCCGCCAAACGATTTTCATGATGTGCTG 147
 120 TTCTGCTCCTGGCTGCAAGATTGCCAATTGATGCCGCCAAACGATTTTCATGATGTGCTG 179
QY 148 GCAATGAAAGACCTTCTGCTTACATGAGGAGACACAATCAATTAATGCTGTCTCTG 207
 180 GCATGAAAGACCTTCTGCTTACATGAGGAGACACAATCAATTAATGCTGTCTCTG 239
QY 208 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGA 267
 240 ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGA 299
Db 268 AAAAATCTCTGGAAGGAGCGCGTGTGCAAGCGGTCTGACCACTGACTCACCAGCCCTCG 327
 300 AAAAATCTCTGGAAGGAGCGCGTGTGCAAGCGGTCTGACCACTGACTCACCAGCCCTCG 359
QY 328 TGGGCTCAAAATATTAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG 387
 360 TGGGCTCAAAATATTAACATTTGCGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG 419
Db 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAATGAGCGTGTATCTGCTGATC 447
 420 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAATGAGCGTGTATCTGCTGATC 479
QY 448 CATATGTTTACAACTGACAGCATGGTCAAGAGCACAGTGAACGGGAAAAATGSCACCGGCC 507
 480 CATATGTTTACAACTGACAGCATGGTCAAGAGCACAGTGAACGGGAAAAATGSCACCGGCC 539
QY 508 AAAGCCATCATACGCTCTTCCCTGATGGGAAACCTTTTCTCAACCAACCCCGATGGAGAA 567
 540 AAAGCCATCATACGCTCTTCCCTGATGGGAAACCTTTTCTCAACCAACCCCGATGGAGAA 599
Db 568 GATGGAATTTCTATCTACGCTCTTCCACACACACTTGGTCAATATTTCAGAAATTGGGACGAT 627
 600 GATGGAATTTCTATCTACGCTCTTCCACACACACTTGGTCAATATTTCAGAAATTGGGACGAT 659
QY 628 GTTCAGTGAGAGTTTCTGTGAACAACGCCAATGTGACACTTGGGCTCAACTCATGGAAG 687
 660 GTTCAGTGAGAGTTTCTGTGAACAACGCCAATGTGACACTTGGGCTCAACTCATGGAAG 719
QY 688 TGACTGTCTACAGAAGACATGACGGGCATATGTTCCCATCGACAAGTGAAGAGATGTGT 747
 720 TGACTGTCTACAGAAGACATGACGGGCATATGTTCCCATCGACAAGTGAAGAGATGTGT 779
Db 748 ACGTGTAAACAGATCAGATTCCTGTGTGTTGTGACTATGTTCCAGAAGAAGATCGAAATT 807
 780 ACGTGTAAACAGATCAGATTCCTGTGTGTTGTGACTATGTTCCAGAAGAAGATCGAAATT 839
QY 808 CATCCGACGAAACCTTCC-CAAGATCTCCCATTTATGTTGATGTCTGATTCATGATC 866
 840 CATCCGACGAAACCTTCTCAAGATCTCCCATTTATGTTGATGTCTGATTCATGATC 899
QY 867 CTAGCCACTTCTCAATTATTTCTACCATTTAACTACAAGTGAGCTTGGGATAAATACTG 926

Db 900 CTAGCCACTTCTCAATTATTTCTACCATTTAACTACAAGTGAGCTTGGGATAAATACTG 959
QY 927 GCCTGTGTTGTTCCACCAATCATCTGTGAATCACACGTAATGTGCTCAATGGAACCTTCA 986
 960 GCCTGTGTTGTTCCACCAATCATCTGTGAATCACACGTAATGTGCTCAATGGAACCTTCA 1019
QY 987 GCCTTAACCTCACTGTGAAGCTGCAGACACGAGACCTTGTCCGCCACCGCACCAACAC 1046
 1020 GCCTTAACCTCACTGTGAAGCTGCAGACACGAGACCTTGTCCGCCACCGCACCAACAC 1079
QY 1047 CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGTGTGACAAACCCCTGAGCTGA 1106
 1080 CCAGACCTTCAAAACCCACCCCTTCTTTAGGACCTGTGTGACAAACCCCTGAGCTGA 1139
QY 1107 GTAGGATTCCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAGGCCACCATCA 1166
 1140 GTAGGATTCCTGATGAAAACTGCCAGATTAAACAGATATGGCCACTTTCAGGCCACCATCA 1199
QY 1167 CAATTGTAGAGGGAATCTTAGAGGTTAACATCATCCAGATGACAGACGTCCTGATGCCGG 1226
 1200 CAATTGTAGAGGGAATCTTAGAGGTTAACATCATCCAGATGACAGACGTCCTGATGCCGG 1259
QY 1227 TGCCATGCGCCTGAAGCTCCCTAATAGACTTTGTGCTGACCTGCAAGGAGCATTCCTCA 1286
 1260 TGCCATGCGCCTGAAGCTCCCTAATAGACTTTGTGCTGACCTGCAAGGAGCATTCCTCA 1319
Db 1287 CGGAGTCTGTACCATCATTTCTGACCCCACTGCGAGATCAACCAAGACACAGCTGTGCA 1346
 1320 CGGAGTCTGTACCATCATTTCTGACCCCACTGCGAGATCAACCAAGACACAGCTGTGCA 1379
QY 1347 GCCCTGTGATGTGATGAGATGTGTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTG 1406
 1380 GCCCTGTGATGTGATGAGATGTGTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTG 1439
QY 1407 GGAAGTACTGTGTGAACCTTCAACCTGCGGGGATGACACAAACCTTGCTCTCACGAGCACCC 1466
 1440 GGAAGTACTGTGTGAACCTTCAACCTGCGGGGATGACACAAACCTTGCTCTCACGAGCACCC 1499
Db 1467 TGAATTTCTGTTCTTGACAGAGACCCAGCCTGCGCTTTAAAGATGGCAAAACAGTGCCTGA 1526
 1500 TGAATTTCTGTTCTTGACAGAGACCCAGCCTGCGCTTTAAAGATGGCAAAACAGTGCCTGA 1559
QY 1527 TCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTTGTGTACAAAAC 1586
 1560 TCTCCGTTGGCTGCTTGGCCATATTTGTCACTGTGATCTCCCTTGTGTACAAAAC 1619
QY 1587 ACAAGGAATACAAACCAATAGAAAATAGTCCCTGGGAATGTGCTCAAGAACAAAGGCTTGA 1646
 1620 ACAAGGAATACAAACCAATAGAAAATAGTCCCTGGGAATGTGCTCAAGAACAAAGGCTTGA 1679
Db 1647 GTGTCTTTCTCAACCGTGCAAAAGCCGCTGTCTTCCCGGAAACAGGAAAGATCCGC 1706
 1680 GTGTCTTTCTCAACCGTGCAAAAGCCGCTGTCTTCCCGGAAACAGGAAAGATCCGC 1739
QY 1707 TACTCAAAAACCAAGAAATTTAAAGAGTCTTCTTAATTTGCACTTGTCTGAAAGCTCA 1766
 1740 TACTCAAAAACCAAGAAATTTAAAGAGTCTTCTTAATTTGCACTTGTCTGAAAGCTCA 1799
QY 1767 CTTTTCAGTGCCATTGATGTGAGATGTGCTGAGTGGCTATTAACTTTTCTTAAG 1826
 1800 CTTTTCAGTGCCATTGATGTGAGATGTGCTGAGTGGCTATTAACTTTTCTTAAG 1859
Db 1827 ATTATGTTAAATAGATATGTTGTTGGGGAAGTTGAATTTTATAGGTTAAATGTC 1886
 1860 ATTATGTTAAATAGATATGTTGTTGGGGAAGTTGAATTTTATAGGTTAAATGTC 1919
QY 1887 TTTTAGAGATGGGAGAGGATTAATACTGCAGGACGCTTCAGCCATGTTGTGAACCTGAT 1946
 1920 TTTTAGAGATGGGAGAGGATTAATACTGCAGGACGCTTCAGCCATGTTGTGAACCTGAT 1979
QY 1947 AAAAGCACTTACGAAAGGCTTTTTCATTAATTTTAATGTTTCACTTATAAGTCTTAG 2006

Db 1980 AAAAGCACTTAGCAAGGCTTCTTTTCATATTTTTTATGTTCACCTATAAAGCTTAG 2039
QY 2007 GTAACTAGTAGATAGAAAACACTGTGTCCCGAGAGTAAGAGAGAGAGTAATTTGATTA 2066
Db 2040 GTAACTAGTAGATAGAAAACACTGTGTCCCGAGAGTAAGAGAGAGAGTAATTTGATTA 2099
QY 2067 GAGCCTAACCCAGGTTAACTGCAAGAAGAGGGGAGTACTTTAGCCTTTCCATGTAAGT 2126
Db 2100 GAGCCTAACCCAGGTTAACTGCAAGAAGAGGGGAGTACTTTAGCCTTTCCATGTAAGT 2159
QY 2127 TATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCA 2186
Db 2160 TATGCATAAAGCCCAATGTAGTCCAGTTTCTAAGATCATGTTCCAAAGCTAACTGAATCCCA 2219
QY 2187 CTTCATATACACACTCATGAAGCTCCTGATGGAACAATAACAGGCCCAAGCCTGTGTATGA 2246
Db 2220 CTTCATATACACACTCATGAAGCTCCTGATGGAACAATAACAGGCCCAAGCCTGTGTATGA 2279
QY 2247 TGTGCACACTTGTCTAGACTCAGAAAAAATACTACTCTCATTAATGGGTGGAGTATTTTG 2306
Db 2280 TGTGCACACTTGTCTAGACTCAGAAAAAATACTACTCTCATTAATGGGTGGAGTATTTTG 2339
QY 2307 GTGACAACTTACTTTGCTTGGCTGAGTGAAGGAATGATATTATATATTATTATTTTCCA 2366
Db 2340 GTGACAACTTACTTTGCTTGGCTGAGTGAAGGAATGATATTATATTATTATTATTTCCA 2399
QY 2367 TGGACATTAGTATGCTTTTATATATACAGGCATGATGCTGAGTGAAGTCTTGTGTA 2426
Db 2400 TGGACATTAGTATGCTTTTATATATACAGGCATGATGCTGAGTGAAGTCTTGTGTA 2459
QY 2427 TATTTCCAAATTTTGTATAGTCGCTGCACATATTTGAATCAAAATATTAAGACTTTCC 2486
Db 2460 TATTTCCAAATTTTGTATAGTCGCTGCACATATTTGAATC-ATATATTAAGACTTTCC 2518
QY 2487 AAAAATTTGTCCTGTTTTTTCATGGCACTTGATCAGTAAGATTCCCTCTGTTTG 2546
Db 2519 AAAGATGAGGTCCCTGTTTTTTCATGGCACTTGATCAGTAAGATTCCCTCTGTTTG 2578
QY 2547 GAACTAAACCATTTACTATATGTTAGACAAGACATTTTTTTTTTCTCTGTAAGAA 2606
Db 2579 TAACTAAACCATTTACTATATGTTAGACATGACATTTCTTCTCTCTGTAAGAA 2638
QY 2607 -AAAAATGAGGAAGAGACAAAAA 2636
Db 2639 TAAAGTGTGGGAAGAGACAAAAA 2669

RESULT 9
US-10-084-817-169
; Sequence 169, Application US/10084817
; Publication No. US20030119009A1
; GENERAL INFORMATION:
; APPLICANT: Susan Stuart
; APPLICANT: Jed G. Nuchtern
; APPLICANT: Sharon E. Plon
; APPLICANT: Jason M. Shohet
; TITLE OF INVENTION: GENES REGULATED BY MYCN ACTIVATION
; FILE REFERENCE: PA-0046 US
; CURRENT APPLICATION NUMBER: US/10/084, 817
; CURRENT FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/270, 784
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 365
; SOFTWARE: PERL Program
; SEQ ID NO 169
; LENGTH: 2666
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030119009A1 2687977CB1
US-10-084-817-169

Query Match 95.1%; Score 2531.6; DB 15; Length 2666;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2576; Conservative 0; Mismatches 19; Indels 3; Gaps 3;
QY 28 AACCTGGTGCCTGCGTCCGTGAGAAATTCACATGGAATGTCTACTATTTCCCTGGAT 87
Db 70 ACCTTGAGTGCCTGCGTCCGTGAGAAATTCACATGGAATGTCTACTATTTCCCTGGAT 129
QY 88 TTCTGCTCTGGCTGCAAGATTGCCCTTGATGCCGCCAAACGATTTCATGATGTGCTGG 147
Db 130 TTCTGCTCTGGCTGCAAGATTGCCCTTGATGCCGCCAAACGATTTCATGATGTGCTGG 189
QY 148 GCAATGAAGAAGCTTCTGCTTACATGAGGAGCACAATCAATTAAATGGCTGCTCTG 207
Db 190 GCAATGAAGAAGCTTCTGCTTACATGAGGAGCACAATCAATTAAATGGCTGCTCTG 249
QY 208 ATGAAAAATGACTGGAATGAAAAAATCTTACCCAGTGTGGAAGCGGGAGACATGAGTGA 267
Db 250 ATGAAAAATGACTGGAATGAAAAAATCTTACCCAGTGTGGAAGCGGGAGACATGAGTGA 309
QY 268 AAAACTCCTGGAAGGAGGCGCGTGTGCAAGCGGCTCTGACAGTGAATCAAGCCCTCG 327
Db 310 AAAACTCCTGGAAGGAGGCGCGTGTGCAAGCGGCTCTGACAGTGAATCAAGCCCTCG 369
QY 328 TGGGCTCAATATTAACATTTGCGGTGAACCTGATATTTCCCTAGATGCCAAAGGAAGATG 387
Db 370 TGGGCTCAATATTAACATTTGCGGTGAACCTGATATTTCCCTAGATGCCAAAGGAAGATG 429
QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGTTATCTGTGATC 447
Db 430 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGTTATCTGTGATC 489
QY 448 CATATGTTTACAACTGACACAGCATGTGTGAGAGACAGTGAACGGGAAATGGACCGGCC 507
Db 490 CGTATGTTTACAACTGACACAGCATGTGTGAGAGACAGTGAACGGGAAATGGACCGGCC 549
QY 508 AAAGCATCATTAACGTCTTCCCTGATGGGAAACCTTTTCTCCACCAACCCCGATGAGAA 567
Db 550 AAAGCATCATTAACGTCTTCCCTGATGGGAAACCTTTTCTCCACCAACCCCGATGAGAA 609
QY 568 GATGAATTTTCATCTACGCTCTTCCACACACTTGGTCAAGTATTTCCAGAAATGGAGCAT 627
Db 610 GATGAATTTTCATCTACGCTCTTCCACACACTTGGTCAAGTATTTCCAGAAATGGAGCAT 669
QY 628 GTTCAGTGAAGTTTCTGTGAACACAGCCCAATGTGACACTTTGGGCTCAACTCATGGAAG 687
Db 670 GTTCAGTGAAGTTTCTGTGAACACAGCCCAATGTGACACTTTGGGCTCAACTCATGGAAG 729
QY 688 TGACTGTCTACAGAGACATGAGCGGGCATATGTTCCCATCGACAAGTGAAGATGTGT 747
Db 730 TGACTGTCTACAGAGACATGAGCGGGCATATGTTCCCATCGACAAGTGAAGATGTGT 789
QY 748 ACGTGTAAACAGATCAGATTCTCTGTGTTGTGACTATGTTCCAGAAAGCATGGAATTT 807
Db 790 ACGTGTAAACAGATCAGATTCTCTGTGTTGTGACTATGTTCCAGAAAGCATGGAATTT 849
QY 808 CATCCGAGGAACCTTCC-CAAGATCTCCCATTAATGTTGATGTCCTGATTCATGATC 866
Db 850 CATCCGAGGAACCTTCTCAAGATCTCCCATTAATGTTGATGTCCTGATTCATGATC 909
QY 867 CTAGCCACTTCTCAATTAATTTCAACATTAACATCAAGTGAAGCTTGGGGATTAATCTG 926
Db 910 CTAGCCACTTCTCAATTAATTTCAACATTAACATCAAGTGAAGCTTGGGGATTAATCTG 969
QY 927 GCCTGTTGTTTCCACCAATCATCTGTGAATCACACGATGTGCTCAATGGAACCTTCA 986
Db 970 GCCTGTTGTTTCCACCAATCATCTGTGAATCACACGATGTGCTCAATGGAACCTTCA 1029
QY 987 GCCTTAACCTCACTGTGAAGCTGCAGACACAGGACCTTGTCCGCCACCGCACCAAC 1046
Db 1030 GCCTTAACCTCACTGTGAAGCTGCAGACACAGGACCTTGTCCGCCACCGCACCAAC 1089
QY 1047 CCAGACCTTCAAAACCAACCCCTTCTTTAGGACCTGTGTGACAAACCCCTGAGCTGA 1106

```
Db 1090 CCAGACCTTCAAAACCCACCCCTTCTTAGGACCTGCTGTGACAAACCCCTGAGACTGA 1149
QY 1107 GTAGATTCTGTATGAAAACCTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCA 1166
Db 1150 GTAGATTCTGTATGAAAACCTGCCAGATTAAACAGATATGGCCACTTTCAAGCCACCATCA 1209
QY 1167 CAATTGTAGAGGAATCTTAGAGGTTAATCATCATCCAGATGACAGACGCTCTGATGCCGG 1226
Db 1210 CAATTGTAGAGGAATCTTAGAGGTTAATCATCATCCAGATGACAGACGCTCTGATGCCGG 1269
QY 1227 TGCCATGGCCCTGAAAAGCTCCCTAATAGACTTTGTCTGACCTGCGCAAGGAGCATTTCCCA 1286
Db 1270 TGCCATGGCCCTGAAAAGCTCCCTAATAGACTTTGTCTGACCTGCGCAAGGAGCATTTCCCA 1329
QY 1287 CGGAGGTCTGTACCATCATTTCTGACCCCACTGCGAGATCAACCCAGAACAGACTCTGCA 1346
Db 1330 CGGAGGTCTGTACCATCATTTCTGACCCCACTGCGAGATCAACCCAGAACAGACTCTGCA 1389
QY 1347 GCCCTGTGAGATGAGATGATGTCTGTGACTGTGAGACGAACCTTCAATGGGTCTG 1406
Db 1390 GCCCTGTGAGATGAGATGATGTCTGTGACTGTGAGACGAACCTTCAATGGGTCTG 1449
QY 1407 GGACGTACTGTGTGAACCTCAACCCCTGGGGGATGACACAAAGCCTGCTCTCACGACACCC 1466
Db 1450 GGACGTACTGTGTGAACCTCAACCCCTGGGGGATGACACAAAGCCTGCTCTCACGACACCC 1509
QY 1467 TGATTTCTGTCTCTGACAGAGACCCAGCCTGCGCTTTAAGGATGGCAACAGTCCCTGA 1526
Db 1510 TGATTTCTGTCTCTGACAGAGACCCAGCCTGCGCTTTAAGGATGGCAACAGTCCCTGA 1569
QY 1527 TCTCCGTTGGCTGCTTGCCCATATTTGTCACTGTGATCTCCCTCTTGCTGTACAAAAAC 1586
Db 1570 TCTCCGTTGGCTGCTTGCCCATATTTGTCACTGTGATCTCCCTCTTGCTGTACAAAAAC 1629
QY 1587 ACAAGGAATACAAACCCATAGAAAAATAGTCTGGGAATGTGTGAGAGCAAAAGGCTGA 1646
Db 1630 ACAAGGAATACAAACCCATAGAAAAATAGTCTGGGAATGTGTGAGAGCAAAAGGCTGA 1689
QY 1647 GTGTCTTTCTCAACCGTGCAAAAGCCGTGTCTTCCCGGAAACAGAAAGATCCGC 1706
Db 1690 GTGTCTTTCTCAACCGTGCAAAAGCCGTGTCTTCCCGGAAACAGAAAGATCCGC 1749
QY 1707 TACTCAAAAACCAAGATTTAAGAGATTCTTAATTTGACCTGTTCTGAAAGCTCA 1766
Db 1750 TACTCAAAAACCAAGATTTAAGAGATTCTTAATTTGACCTGTTCTGAAAGCTCA 1809
QY 1767 CTTTTCAGTGCCATTGATGTGAGATGTCTGAGTGGCTATTAACTTTTTCCTAAG 1826
Db 1810 CTTTTCAGTGCCATTGATGTGAGATGTCTGAGTGGCTATTAACTTTTTCCTAAG 1869
QY 1827 ATTATTGTTAATAGATATTGTGTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA 1886
Db 1870 ATTATTGTTAATAGATATTGTGTTGGGGAAGTTGAATTTTATAGGTTAAATGTCA 1929
QY 1887 TTTTAGAGATGGGAGAGGAGATTATACTGACAGCAGCTTCAGCCATGTTGAAAAGCTGAT 1946
Db 1930 TTTTAGAGATGGGAGAGGAGATTATACTGACAGCAGCTTCAGCCATGTTGAAAAGCTGAT 1989
QY 1947 AAAAGCAACTTAGCAAGGCTCTTTTCATTATTTTATATGTTTCACTTAAGAGCTTAG 2006
Db 1990 AAAAGCAACTTAGCAAGGCTCTTTTCATTATTTTATATGTTTCACTTAAGAGCTTAG 2049
QY 2007 GTAAGTAGTAGATAGAAACACATGTCTCCGAGAGTAAGAGAGAAAGCTACTATTGATTA 2066
Db 2050 GTAAGTAGTAGATAGAAACACATGTCTCTCCGAGAGTAAGAGAGAAAGCTACTATTGATTA 2109
QY 2067 GAGCCTAACCCAGGTTAACTGTGAGAGAGAGGGCGGATACCTTTTCCATGTAAGTGTG 2126
Db 2110 GAGCCTAACCCAGGTTAACTGTGAGAGAGAGGGCGGATACCTTTTCCATGTAAGTGTG 2169
QY 2127 TATGATAAAGCCAAATGTAAGTGTGCTTAAGATCATGTTCAGCTAACTGAATCCCA 2186
```

```
Db 2170 TATGATAAAGCCAAATGTAAGTGTGCTCCAGTTTCTAAGATCATGTTCACAGCTAACTGAATCCCA 2229
QY 2187 CTTCAATACACACTCATGAAGTCTCCTGATGGAACAATAACAGGCCCAAGCCTGTGTATGA 2246
Db 2230 CTTCAATACACACTCATGAAGTCTCCTGATGGAACAATAACAGGCCCAAGCCTGTGTATGA 2289
QY 2247 TGTGCACACTTGTGTAGACTCAGAAAAAATACTACTCTCATTAATGGGTGGAGTATTTTG 2306
Db 2290 TGTGCACACTTGTGTAGACTCAGAAAAAATACTACTCTCATTAATGGGTGGAGTATTTTG 2349
QY 2307 GTGACAACCTACTTTGCTTGCTGGCTGAGTGAAGGAATGATATTATATATCAATTTATTTCCA 2366
Db 2350 GTGACAACCTACTTTGCTTGCTGGCTGAGTGAAGGAATGATATTATATATCAATTTATTTCCA 2409
QY 2367 TGGACATTTAGTTAGTGTCTTTTATATACCAAGCATGATGCTGAGTGAACCTCTGTGTA 2426
Db 2410 TGGACATTTAGTTAGTGTCTTTTATATACCAAGCATGATGCTGAGTGAACCTCTGTGTA 2469
QY 2427 TATTTCCAAATTTTGTATAGTGTGCTGCACATATTTGAATCAAAATATTAAGACTTTCC 2486
Db 2470 TATTTCCAAATTTTGTATAGTGTGCTGCACATATTTGAATC-ATATATTAAGACTTTCC 2528
QY 2487 AAAAATTTGCTCCCTGCTGTTTTCATGCACTTGATCATGTAAGGATTTCCCTGTGTTG 2546
Db 2529 AAAGATGAGTCTCCCTGCTGTTTTCATGCACTTGATCATGTAAGGATTTCACTCTGTGTTG 2588
QY 2547 GAACCTAAACCATTTACTATATGTTAGACAAGACATTTTCTTCTCTCTGAAAAA 2606
Db 2589 TAACTAATAACCATCTACTATATGTTAGACATGACATTCCTTTCTCTCTCTGAAAAA 2648
QY 2607 -AAAATGAGGGAAGAGAC 2623
Db 2649 TAAAGTGTGGGAAGAGAC 2666

RESULT 10
US-10-219-535-41
; Sequence 41, Application US/10219535
; Publication No. US20040044179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C60
; CURRENT APPLICATION NUMBER: US/10/219,535
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
```

; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 41
; LENGTH: 2683
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-535-41

Query Match 93.4%; Score 2485; DB 13; Length 2683;
Best Local Similarity 97.8%; Pred. No. 0;
Matches 2576; Conservative 0; Mismatches 20; Indels 39; Gaps 4;

QY 28 AACCTTGTCCTGCGTCCGTGAGAAATCAGCATGGAATGCTCTACTATTCTCTGGAT 87
Db 50 ACCTTGAGTGCCTGCGTCCGTGAGAAATCAGCATGGAATGCTCTACTATTCTCTGGAT 109
QY 88 TTCTGCTCCTGCGTGCAGAGATTGCCACTTGATGCCGCCAAACGATTTCATGATGTGCTGG 147
Db 110 TTCTGCTCCTGCGTGCAGAGATTGCCACTTGATGCCGCCAAACGATTTCATGATGTGCTGG 169
QY 148 GCAATGAAGAACCCTTCTGCTTACATGAGGAGCACAAATCAATTAATGCGCTGCTCTG 207
Db 170 GCAATGAAGAACCCTTCTGCTTACATGAGGAGCACAAATCAATTAATGCGCTGCTCTG 229
QY 208 ATGAAAATGACTGGAAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGA 267
Db 230 ATGAAAATGACTGGAAATGAAAACTCTACCCAGTGTGGAAGCGGGAGACATGAGGTGA 289
QY 268 AAAACTCCTGGAAGGAGCGCGTGTGCAGGCGGCTTACCAAGTGACTCACCGCCCTCG 327
Db 290 AAAACTCCTGGAAGGAGCGCGTGTGCAGGCGGCTTACCAAGTGACTCACCGCCCTCG 349
QY 328 TGGGCTCAAAATATAACATTGCGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG 387
Db 350 TGGGCTCAAAATATAACATTGCGGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG 409
QY 388 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGTTTACTGCTGATC 447
Db 410 CCAATGGCAACATAGTCTATGAGAAGAACTGCAGAAATGAGGCTGTTTACTGCTGATC 469
QY 448 CATATGTTTACAACGTGACAGCATGGTCAAGAGACGTGACGGGAAATGGCACCGGCC 507
Db 470 CGTATGTTTACAACGTGACAGCATGGTCAAGAGACGTGACGGGAAATGGCACCGGCC 529
QY 508 AAAGCCATGATAACGTTCTCCCTGATGGGAAACCTTTTCTCAACCAACCCCGGATGGAAG 567
Db 530 AAAGCCATGATAACGTTCTCCCTGATGGGAAACCTTTTCTCAACCAACCCCGGATGGAAG 589
QY 568 GATGGAATTTCACTACGTTCTCCACACACACTTGGTCAATTTCCAGAAATTGGGACGAT 627
Db 590 GATGGAATTTCACTACGTTCTCCACACACACTTGGTCAATTTCCAGAAATTGGGACGAT 649
QY 628 GTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG 687
Db 650 GTTCAGTGAGAGTTTCTGTGAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG 709
QY 688 TGACTGTCTACAGAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 747
Db 710 TGACTGTCTACAGAGACATGGACGGGCATATGTTCCCATCGCACAAAGTGAAGATGTGT 769
QY 748 ACGTGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCAGAGAAGCATCGAAATT 807
Db 770 ACGTGTAAACAGATCAGATTCCTGTGTTGTGACTATGTTCAGAGAAGCATCGAAATT 829
QY 808 CATCCGACGAACCTTCC-CAGAGATCTCCCATTTATGTTGATGTCCTGATTCAATGATC 866
Db 830 CATCCGACGAACCTTCCCTCAGAGATCTCCCATTTATGTTGATGTCCTGATTCAATGATC 889
QY 867 CTAGCCACTTCTCAATTATTCTAACCAATTACTACAAGTGAAGCTTCGGGATATAACTG 926

Db 890 CTAGCCACTTCTCAATTATTCTAACCAATTACTACAAGTGAAGCTTCGGGATATAACTG 949
QY 927 GCCTGTGTTGTTCCACCAATCATCTGTGAATCACACGTATGTCTCAATGGAACCTTCA 986
Db 950 GCCTGTGTTGTTCCACCAATCATCTGTGAATCACACGTATGTCTCAATGGAACCTTCA 1009
QY 987 GCCTTAACCTCACTGTGAAGCTGCAGACACCAGACCTTGTCCGCCACCGCACCAACCAC 1046
Db 1010 GCCTTAACCTCACTGTGAAGCTGCAGACACCAGACCTTGTGTCCGCCACCGCACCAACCAC 1069
QY 1047 CCAGACCTTCAAAACCACCCCTTCTTT----- 1074
Db 1070 CCAGACCTTCAAAACCACCCCTTCTTTAGCAACTACTCTAAATCTTATGATTCAACACA 1129
QY 1075 ----AGGACTGCTGTGTGACCAACCCCTGAGAGCTGAGTAGATTCCTGATGAACCTGCC 1130
Db 1130 CCCAGGACCTACTGTGTGACCAACCCCTGAGAGCTGAGTAGATTCCTGATGAACCTGCC 1189
QY 1131 AGATTAAAGATAATGGCCACTTTCAAGCCACCATCAATTTGAGAGGAATCTTAGAGG 1190
Db 1190 AGATTAAAGATAATGGCCACTTTCAAGCCACCATCAATTTGAGAGGAATCTTAGAGG 1249
QY 1191 TTAACATCATCCAGATGACAGACGTCCTGATGCGGTGCCATGGCTGAAAGCTCCCTAA 1250
Db 1250 TTAACATCATCCAGATGACAGACGTCCTGATGCGGTGCCATGGCTGAAAGCTCCCTAA 1309
QY 1251 TAGACTTTGTGTGTAACCTGCCAAGGAGCATTCCCAACGAGGTCTGTACATCATTTCTG 1310
Db 1310 TAGACTTTGTGTGTAACCTGCCAAGGAGCATTCCCAACGAGGTCTGTACATCATTTCTG 1369
QY 1311 ACCCCACCTGGAGATCAACCCAGAACACAGTCTGCAGCCCTGTGGATGTGATGAGATGT 1370
Db 1370 ACCCCACCTGGAGATCAACCCAGAACACAGTCTGCAGCCCTGTGGATGTGATGAGATGT 1429
QY 1371 GTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGAGCGTACTGTGTGAACCTCAACC 1430
Db 1430 GTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGAGCGTACTGTGTGAACCTCAACC 1489
QY 1431 TGGGGATGACAAAGCCTGGCTCTACAGACACCCTGATTTCTGTTCTGACAGAGACC 1490
Db 1490 TGGGGATGACAAAGCCTGGCTCTACAGACACCCTGATTTCTGTTCTGACAGAGACC 1549
QY 1491 CAGCCTCGCCTTTAAGATGGCAAAACAGTGCCCTGATCTCCGTTGAGCTGTGGCCATAT 1550
Db 1550 CAGCCTCGCCTTTAAGATGGCAAAACAGTGCCCTGATCTCCGTTGAGCTGTGGCCATAT 1609
QY 1551 TTGTCACTGTGATCTCCCTCTTGTGTACAAAAAACAACAAGGATACAAACCCCATAGAAA 1610
Db 1610 TTGTCACTGTGATCTCCCTCTTGTGTACAAAAAACAACAAGGATACAAACCCCATAGAAA 1669
QY 1611 ATAGTCCGTGGAAATGTGTCAGAAAGCAAAAGGCTGAGTGTCTTCTCAACCCGTGCAAAAG 1670
Db 1670 ATAGTCCGTGGAAATGTGTCAGAAAGCAAAAGGCTGAGTGTCTTCTCAACCCGTGCAAAAG 1729
QY 1671 CCGTGTCTTCCCGGGAACAAGAAAAGGATCCGCTACTCAAAAACCAAGAAATTAAAG 1730
Db 1730 CCGTGTCTTCCCGGGAACAAGAAAAGGATCCGCTACTCAAAAACCAAGAAATTAAAG 1789
QY 1731 GAGTTTCTTAATTTGCACTTGTCTTCTGAAGCTCACTTTTCAAGTGCATTTGATGAGAG 1790
Db 1790 GAGTTTCTTAATTTGCACTTGTCTTCTGAAGCTCACTTTTCAAGTGCATTTGATGAGAG 1849
QY 1791 TGTGCTGAGATGGCTATTAACCTTTTTTCTCTAAAGATTATGTCTTAATAGATATTGTGG 1850
Db 1850 TGTGCTGAGATGGCTATTAACCTTTTTTCTCTAAAGATTATGTCTTAATAGATATTGTGG 1909
QY 1851 TTTGGGGAAGTTGAATTTTTTATAGGTTAAATGTCAATTTTAGAGATGGGAGAGGAAATTA 1910
Db 1910 TTTGGGGAAGTTGAATTTTTTATAGGTTAAATGTCAATTTTAGAGATGGGAGAGGAAATTA 1969
QY 1911 TACTGACAGCAGCTTCAGCCATGTTGTGAACCTGATMAAAGCAACTTAGCAAGGCTTCTT 1970
Db 1970 TACTGACAGCAGCTTCAGCCATGTTGTGAACCTGATMAAAGCAACTTAGCAAGGCTTCTT 2029

PRIOR FILING DATE: 1998-08-11
 PRIOR APPLICATION NUMBER: 60/096791
 PRIOR FILING DATE: 1998-08-17
 PRIOR APPLICATION NUMBER: 60/097986
 PRIOR FILING DATE: 1998-08-26
 PRIOR APPLICATION NUMBER: 60/098544
 PRIOR FILING DATE: 1998-08-31
 PRIOR APPLICATION NUMBER: 60/099596
 PRIOR FILING DATE: 1998-09-09
 PRIOR APPLICATION NUMBER: 60/099598
 PRIOR FILING DATE: 1998-09-09
 PRIOR APPLICATION NUMBER: 60/099803
 PRIOR FILING DATE: 1998-09-10
 PRIOR APPLICATION NUMBER: 60/099811
 PRIOR FILING DATE: 1998-09-10
 PRIOR APPLICATION NUMBER: 60/099812
 PRIOR FILING DATE: 1998-09-10
 PRIOR APPLICATION NUMBER: 60/099816
 PRIOR FILING DATE: 1998-09-10
 PRIOR APPLICATION NUMBER: 60/100038
 PRIOR FILING DATE: 1998-09-11
 PRIOR APPLICATION NUMBER: 60/100385
 PRIOR FILING DATE: 1998-09-15
 PRIOR APPLICATION NUMBER: 60/100390
 PRIOR FILING DATE: 1998-09-15
 PRIOR APPLICATION NUMBER: 60/100627
 PRIOR FILING DATE: 1998-09-16
 PRIOR APPLICATION NUMBER: 60/100848
 PRIOR FILING DATE: 1998-09-18
 PRIOR APPLICATION NUMBER: 60/100919
 PRIOR FILING DATE: 1998-09-17
 PRIOR APPLICATION NUMBER: 60/101477
 PRIOR FILING DATE: 1998-09-23
 PRIOR APPLICATION NUMBER: 60/101738
 PRIOR FILING DATE: 1998-09-24
 PRIOR APPLICATION NUMBER: 60/101741
 PRIOR FILING DATE: 1998-09-24
 PRIOR APPLICATION NUMBER: 60/101786
 PRIOR FILING DATE: 1998-09-25
 PRIOR APPLICATION NUMBER: 60/101916
 PRIOR FILING DATE: 1998-09-24
 PRIOR APPLICATION NUMBER: 60/101922
 PRIOR FILING DATE: 1998-09-24
 PRIOR APPLICATION NUMBER: 60/106178
 PRIOR FILING DATE: 1998-10-28
 PRIOR APPLICATION NUMBER: 60/106248
 PRIOR FILING DATE: 1998-10-29
 PRIOR APPLICATION NUMBER: 60/106464
 PRIOR FILING DATE: 1998-10-30
 PRIOR APPLICATION NUMBER: 60/106905
 PRIOR FILING DATE: 1998-11-03
 PRIOR APPLICATION NUMBER: 60/108787
 PRIOR FILING DATE: 1998-11-17
 PRIOR APPLICATION NUMBER: 60/108801
 PRIOR FILING DATE: 1998-11-17
 PRIOR APPLICATION NUMBER: 60/108849
 PRIOR FILING DATE: 1998-11-18
 PRIOR APPLICATION NUMBER: 60/112422
 PRIOR FILING DATE: 1998-12-15
 PRIOR APPLICATION NUMBER: 60/113296
 PRIOR FILING DATE: 1998-12-22
 PRIOR APPLICATION NUMBER: 60/113605
 PRIOR FILING DATE: 1998-12-23
 PRIOR APPLICATION NUMBER: 60/113621
 PRIOR FILING DATE: 1998-12-23
 PRIOR APPLICATION NUMBER: 60/115558
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/115565
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/115733
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/119549
 PRIOR FILING DATE: 1999-02-10

PRIOR APPLICATION NUMBER: 60/123618
 PRIOR FILING DATE: 1999-03-10
 PRIOR APPLICATION NUMBER: 60/125259
 PRIOR FILING DATE: 1999-03-19
 PRIOR APPLICATION NUMBER: 60/125775
 PRIOR FILING DATE: 1999-03-23
 PRIOR APPLICATION NUMBER: 60/126773
 PRIOR FILING DATE: 1999-03-29
 PRIOR APPLICATION NUMBER: 60/127887
 PRIOR FILING DATE: 1999-04-05
 PRIOR APPLICATION NUMBER: 60/130232
 PRIOR FILING DATE: 1999-04-21
 PRIOR APPLICATION NUMBER: 60/131022
 PRIOR FILING DATE: 1999-04-26
 PRIOR APPLICATION NUMBER: 60/131270
 PRIOR FILING DATE: 1999-04-27
 PRIOR APPLICATION NUMBER: 60/131291
 PRIOR FILING DATE: 1999-04-27
 PRIOR APPLICATION NUMBER: 60/131445
 PRIOR FILING DATE: 1999-04-28
 PRIOR APPLICATION NUMBER: 60/134287
 PRIOR FILING DATE: 1999-05-14
 PRIOR APPLICATION NUMBER: 60/140650
 PRIOR FILING DATE: 1999-06-22
 PRIOR APPLICATION NUMBER: 60/140723
 PRIOR FILING DATE: 1999-06-22
 PRIOR APPLICATION NUMBER: 60/141037
 PRIOR FILING DATE: 1999-06-23
 PRIOR APPLICATION NUMBER: 60/144758
 PRIOR FILING DATE: 1999-07-20
 PRIOR APPLICATION NUMBER: 60/145698
 PRIOR FILING DATE: 1999-07-26
 PRIOR APPLICATION NUMBER: 60/146222
 PRIOR FILING DATE: 1999-07-28
 PRIOR APPLICATION NUMBER: 60/146963
 PRIOR FILING DATE: 1999-08-03
 PRIOR APPLICATION NUMBER: 60/149320
 PRIOR FILING DATE: 1999-08-17
 PRIOR APPLICATION NUMBER: 60/149638
 PRIOR FILING DATE: 1999-08-17
 PRIOR APPLICATION NUMBER: 60/151733
 PRIOR FILING DATE: 1999-08-31
 PRIOR APPLICATION NUMBER: 60/164418
 PRIOR FILING DATE: 1999-11-09
 PRIOR APPLICATION NUMBER: 60/166361
 PRIOR FILING DATE: 1999-11-16
 PRIOR APPLICATION NUMBER: 60/169445
 PRIOR FILING DATE: 1999-12-07
 PRIOR APPLICATION NUMBER: 60/169495
 PRIOR FILING DATE: 1999-12-07
 PRIOR APPLICATION NUMBER: 60/169835

Query Match 93.4%; Score 2485; DB 15; Length 2683;
 Best Local Similarity 97.8%; Pred. No. 0;
 Matches 2576; Conservative 0; Mismatches 20; Indels 39; Gaps 4;

QY 28 AACCTTGCTGCTGCGTCCGTGAGAAATTCAGCAATGATGCTCTACTATTTCCTGGAT 87
 Db 50 ACCTTGAGTGCCTGCGTCCGTGAGAAATTCAGCAATGATGCTCTACTATTTCCTGGAT 109
 QY 88 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCCAACGATTTTCATGATGTGCTGG 147
 Db 110 TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCCAACGATTTTCATGATGTGCTGG 169
 QY 148 GCAATGAAGACCTTCTGCTTACATGAGGAGCAATCAATTAATGCTGCTCTG 207
 Db 170 GCAATGAAGACCTTCTGCTTACATGAGGAGCAATCAATTAATGCTGCTCTG 229
 QY 208 ATGAAATGACTGGAATGAAAACTTACCCAGTGTGGAAGCGGAGACATGAGTGA 267
 Db 230 ATGAAATGACTGGAATGAAAACTTACCCAGTGTGGAAGCGGAGACATGAGTGA 289
 QY 268 AAAACTCCTGGAAGGAGCGCGTGTGACAGCGGTCTTGACCAAGTGACTACAGCCCTCG 327

Db 290 AAAAATCCTGGAAGGAGGCGCTGTGCAGGCGGTCCCTGACCACTGACTCACCAAGCCCTCG 349
QY 328 TGGGCTCAATATTAACATTTTGGCGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG 387
Db 350 TGGGCTCAATATTAACATTTTGGCGTGAACCTGATATTCCTAGATGCCAAAAGGAAGATG 409
QY 388 CCAATGGCAACATAGTCTATGAGAAACTGCAGAAATGAGGCTGTTTATCTGCTGATC 447
Db 410 CCAATGGCAACATAGTCTATGAGAAACTGCAGAAATGAGGCTGTTTATCTGCTGATC 469
QY 448 CATATGTTTACAACCTGGACAGCATGGTCAAGAGACAGTGAACGGGAAAAATGGCACCGGCC 507
Db 470 CGTATGTTTACAACCTGGACAGCATGGTCAAGAGACAGTGAACGGGAAAAATGGCACCGGCC 529
QY 508 AAAGCCATCATACGTCTTCCCTGATGGGAAACCTTTTCCCTCACCAACCCCGGATGAGAA 567
Db 530 AAAGCCATCATACGTCTTCCCTGATGGGAAACCTTTTCCCTCACCAACCCCGGATGAGAA 589
QY 568 GATGGAATTTCACTACGTCTTCCACACACTTGGTCAATTTTCCAGAAATTGGAGCAT 627
Db 590 GATGGAATTTCACTACGTCTTCCACACACTTGGTCAATTTTCCAGAAATTGGAGCAT 649
QY 628 GTTCAGTGAAGTTTCTGTGAACACAGCCATGTGACACTTGGGCTTCAACTCATGGAAG 687
Db 650 GTTCAGTGAAGTTTCTGTGAACACAGCCAAATGTGACACTTGGGCTTCAACTCATGGAAG 709
QY 688 TGAATGTTTACAGAAAGACATGGAACGGGCATATGTCCCATGCGACAAAGTGAAGATGTGT 747
Db 710 TGAATGTTTACAGAAAGACATGGAACGGGCATATGTCCCATGCGACAAAGTGAAGATGTGT 769
QY 748 ACGTGTACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAAGCATCGAAATT 807
Db 770 ACGTGTACAGATCAGATTCCTGTGTTGTGACTATGTTCCAGAAAGCATCGAAATT 829
QY 808 CATCCGACGAAACCTTCC-CAAAGATCTCCCAATTATGTTGATGTCTGATTCATGATC 866
Db 830 CATCCGACGAAACCTTCCCAAAGATCTCCCAATTATGTTGATGTCTGATTCATGATC 889
QY 867 CTAGCCACTTCTCAATTAATTCTAACCATTAATCTACAAAGTGAAGCTTCGGGGAATAACTG 926
Db 890 CTAGCCACTTCTCAATTAATTCTAACCATTAATCTACAAAGTGAAGCTTCGGGGAATAACTG 949
QY 927 GCCTGTTGTTTCCACCAATCATACTGTGAATCACACGTATGTCTCAATGGAACCTTCA 986
Db 950 GCCTGTTGTTTCCACCAATCATACTGTGAATCACACGTATGTCTCAATGGAACCTTCA 1009
QY 987 GCCTTAACCTCACTGTGAAGCTGCAGACACGAGACCTTGTCCGCCACCGCCACCAACAC 1046
Db 1010 GCCTTAACCTCACTGTGAAGCTGCAGACACGAGACCTTGTCCGCCACCGCCACCAACAC 1069
QY 1047 CCAAGACCTTCAAAACCCACCCCTTCTT----- 1074
Db 1070 CCAAGACCTTCAAAACCCACCCCTTCTTGAACACTACTCTAAATCTTATGATTCAAACA 1129
QY 1075 ----AGGAAGCTGCTGTGACAAACCCCTGAGCTGAGTAGAATTCCTGATGAAAACTGCC 1130
Db 1130 CCCCAGGAACTTACTGTGACAAACCCCTGAGCTGAGTAGAATTCCTGATGAAAACTGCC 1189
QY 1131 AGATTAACAGATATGGCCACTTTTCAAGCCCAACCATCACAATTTGTAGAGGGAATCTTAGAG 1190
Db 1190 AGATTAACAGATATGGCCACTTTTCAAGCCCAACCATCACAATTTGTAGAGGGAATCTTAGAG 1249
QY 1191 TTAACATCATCCAGATGACAGACGCTCTGATGCCGCTGCATGAGCTGAAAGCTCCCTAA 1250
Db 1250 TTAACATCATCCAGATGACAGACGCTCTGATGCCGCTGCATGAGCTGAAAGCTCCCTAA 1309
QY 1251 TAGACTTTGTGCTGACCTGCAAGGAGCATTCCTCAAGAGGTCTGTACCATCATTTCTG 1310
Db 1310 TAGACTTTGTGCTGACCTGCAAGGAGCATTCCTCAAGAGGTCTGTACCATCATTTCTG 1369
QY 1311 ACCCCACCTGCGAGATCAACCCAGAACACAGTCTGACGCCCTGTGATGTGATGAGATGT 1370

Db 1370 ACCCCACCTGCGAGATCAACCCAGAACACAGTCTGCAAGCCCTGTGATGTGATGAGATGT 1429
QY 1371 GTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCC 1430
Db 1430 GTCTGCTGACTGTGAGACGAACCTTCAATGGGTCTGGGACGTACTGTGTGAACCTCACCC 1489
QY 1431 TGGGGATGACACAAGCCTGCTCTCACGAGCACCTGATTTCTGTCTCTGACAGAGACC 1490
Db 1490 TGGGGATGACACAAGCCTGCTCTCACGAGCACCTGATTTCTGTCTCTGACAGAGACC 1549
QY 1491 CAGCCTCGCCTTTAAGGATGGCAAAACAGTGGCCCTGATCTCCGTTGGCTGCTTGGCCATAT 1550
Db 1550 CAGCCTCGCCTTTAAGGATGGCAAAACAGTGGCCCTGATCTCCGTTGGCTGCTTGGCCATAT 1609
QY 1551 TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACAACAAGGAATACAACCCATAGAAA 1610
Db 1610 TTGTCACTGTGATCTCCCTCTTGGTGTACAAAAACAACAAGGAATACAACCCATAGAAA 1669
QY 1611 ATAGTCTGGGAATGTGTCAAGAACAAAGCCCTGAGTGTCTTCTCAACCGTGCAAAAG 1670
Db 1670 ATAGTCTGGGAATGTGTCAAGAACAAAGCCCTGAGTGTCTTCTCAACCGTGCAAAAG 1729
QY 1671 CCGTGTCTTCCCGGAAACCAAGAAAGATCCGCTACTCAAAAACAAGAAATTTAAAG 1730
Db 1730 CCGTGTCTTCCCGGAAACCAAGAAAGATCCGCTACTCAAAAACAAGAAATTTAAAG 1789
QY 1731 GAGTTTCTTAAATTTTGAACCTTGTCTGAAGCTCACTTTTCAAGTGCATGTGAGGA 1790
Db 1790 GAGTTTCTTAAATTTTGAACCTTGTCTGAAGCTCACTTTTCAAGTGCATGTGAGGA 1849
QY 1791 TGTGCTGAGTGCTATTAAACCTTTTTCCTAAAGATTATTTGTTAAATAGATATTGTGG 1850
Db 1850 TGTGCTGAGTGCTATTAAACCTTTTTCCTAAAGATTATTTGTTAAATAGATATTGTGG 1909
QY 1851 TTTGGGAAGTTGAATTTTATAGTTAAATGTCAATTTTAGAGATGGGAGAGGATTA 1910
Db 1910 TTTGGGAAGTTGAATTTTATAGTTAAATGTCAATTTTAGAGATGGGAGAGGATTA 1969
QY 1911 TACTGACGAGCTTCAGCCATGTTGTGAACCTGATAAAGCAACCTTAGCAAGGCTTCTT 1970
Db 1970 TACTGACGAGCTTCAGCCATGTTGTGAACCTGATAAAGCAACCTTAGCAAGGCTTCTT 2029
QY 1971 TTCAATTAATTTTATGTTTCACTTATAAAGCTTAGGTAACCTAGTAGATAGAAACACTG 2030
Db 2030 TTCAATTAATTTTATGTTTCACTTATAAAGCTTAGGTAACCTAGTAGATAGAAACACTG 2089
QY 2031 TGTCCGAGAGTAAAGAGAGAAAGCTACTATTGATTAGAGCTTAACCCAGGTTAACTGCA 2090
Db 2090 TGTCCGAGAGTAAAGAGAGAAAGCTACTATTGATTAGAGCTTAACCCAGGTTAACTGCA 2149
QY 2091 GAAAGAGCGGATACTTTCAGCTTTCATGTGAATCTGTATGCAATAAAGCCATGTAGTCA 2150
Db 2150 GAAAGAGCGGATACTTTCAGCTTTCATGTGAATCTGTATGCAATAAAGCCATGTAGTCA 2209
QY 2151 GTTCTTAAGATCATGTTTCAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTCC 2210
Db 2210 GTTCTTAAGATCATGTTTCAAGCTAACTGAATCCCACTTCAATACACACTCATGAACCTCC 2269
QY 2211 TGATGGAACAATAACAGGCCCAAGCCTGTGTATGATGTCACACTTGTGACTCAGAA 2270
Db 2270 TGATGGAACAATAACAGGCCCAAGCCTGTGTATGATGTCACACTTGTGACTCAGAA 2329
QY 2271 AAATACTACTCTCATAAATGGGTGGAGATATTTTGGTGAACAACCTACTTGTGGCTG 2330
Db 2330 AAATACTACTCTCATAAATGGGTGGAGATATTTTGGTGAACAACCTACTTGTGGCTG 2389
QY 2331 AGTGAAGGAATGATATTTCATATATTCATTTATTCATGAGCATTTAGTGTCTTTTA 2390
Db 2390 AGTGAAGGAATGATATTTCATATATTCATTTATTCATGAGCATTTAGTGTCTTTTA 2449
QY 2391 TATACAGGCATGATGTCTGAGTGACACTTGTGTATATTTCCAAATTTTGTATAGTCTG 2450
Db 2450 TATACAGGCATGATGTCTGAGTGACACTTGTGTATATTTCCAAATTTTGTATAGTCTG 2509

QY	2451	CTGCACATATTGTAATCAAAATATTAGACTTTCCAAAATTGGTCCCCTGTTTTTCA	2510
Db	2510	CTGCACATATTGAATC-ATATATTAGACTTTCCAAAGATGAGGTCCCTGTTTTTCA	2568
QY	2511	TGGCAACTTGATCAGTAGAATTTCCCTCTGTTTGAACATAAACCATTTACTATATGT	2570
Db	2569	TGGCAACTTGATCAGTAGAATTTCACTCTGTTTGTACTAAAAACCATCTACTATATGT	2628
QY	2571	TAGACAAGACATTTTTTTTTTTCCTTCCCTGA AAAA-AAATGAGGGAAGAGACA	2624
Db	2629	TAGACATGACATTTCTTTCTCTCCTCTCGAAAAATAAAGTGTGGGAAGAGACA	2683

RESULT 14

US-10-230-163-41

Sequence 41, Application US/10230163
Publication No. US20030036635A1

```

: APPLICANT: Baker, Kevin P.
: APPLICANT: Desnoyers, Luc
: APPLICANT: Gerritsen, Mary
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Grimaldi, J. Christopher
: APPLICANT: Gurney, Austin L.
: APPLICANT: Smith, Victoria
: APPLICANT: Stephan, Jean-Philippe F.
: APPLICANT: Watanabe, Colin L.
: APPLICANT: Wood, William I.
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: ACIDS ENCODING THE SAME
: FILE REFERENCE: P3530P1C96

```

QY	2451	CTGCACATATTTGAATCAAAATATTAAGACTTTCCAAAATTTGGTCCCTGGTTTTC	2510
Db	2510	CTGCACATATTTGAATC-ATATATTAAGACTTTCCAAAGATGAGGTCCCTGGTTTTC	2568
QY	2511	TGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTGAACTAAACCATTACTATATGT	2570
Db	2569	TGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTGTAATAAACCATCTACTATATGT	2628
QY	2571	TAGACAAGACATTTTTTTTTTCTCCTCGAAAA-AAAATGAGGGAAGAGACA	2624
Db	2629	TAGACATGACATCTTTTCTCTCCTCTGAAAAATAAGTGTGGGAAGAGACA	2683
RESULT 14			
US-10-230-163-41			
Sequence 41, Application US/10230163			
Publication No. US20030036635A1			
GENERAL INFORMATION:			
APPLICANT: Baker, Kevin P.			
APPLICANT: Desnoyers, Luc			
APPLICANT: Gerlitsen, Mary			
APPLICANT: Goddard, Audrey			
APPLICANT: Godowski, Paul J.			
APPLICANT: Grimaldi, J. Christopher			
APPLICANT: Gurney, Austin L.			
APPLICANT: Smith, Victoria			
APPLICANT: Stephan, Jean-Philippe F.			
APPLICANT: Watanabe, Colin L.			
APPLICANT: Wood, William I.			
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC			
FILE REFERENCE: P3530P1C96			
CURRENT APPLICATION NUMBER: US/10/230,163			
CURRENT FILING DATE: 2002-08-28			
PRIOR APPLICATION NUMBER: 10/119,480			
PRIOR FILING DATE: 2002-04-09			
PRIOR APPLICATION NUMBER: 60/059113			
PRIOR FILING DATE: 1997-09-17			
PRIOR APPLICATION NUMBER: 60/062287			
PRIOR FILING DATE: 1997-10-17			
PRIOR APPLICATION NUMBER: 60/063549			
PRIOR FILING DATE: 1997-10-28			
PRIOR APPLICATION NUMBER: 60/064103			
PRIOR FILING DATE: 1997-10-31			
PRIOR APPLICATION NUMBER: 60/069873			
PRIOR FILING DATE: 1997-12-17			
PRIOR APPLICATION NUMBER: 60/078910			
PRIOR FILING DATE: 1998-03-20			
PRIOR APPLICATION NUMBER: 60/079294			
PRIOR FILING DATE: 1998-03-25			
PRIOR APPLICATION NUMBER: 60/079656			
PRIOR FILING DATE: 1998-03-26			
PRIOR APPLICATION NUMBER: 60/079728			
PRIOR FILING DATE: 1998-03-27			
PRIOR APPLICATION NUMBER: 60/081819			
PRIOR FILING DATE: 1998-04-15			
PRIOR APPLICATION NUMBER: 60/081955			
PRIOR FILING DATE: 1998-04-15			
PRIOR APPLICATION NUMBER: 60/082804			
PRIOR FILING DATE: 1998-04-22			
PRIOR APPLICATION NUMBER: 60/084441			
PRIOR FILING DATE: 1998-05-06			
PRIOR APPLICATION NUMBER: 60/085323			
PRIOR FILING DATE: 1998-05-13			
PRIOR APPLICATION NUMBER: 60/085579			
PRIOR FILING DATE: 1998-05-15			
PRIOR APPLICATION NUMBER: 60/086392			
PRIOR FILING DATE: 1998-05-22			
PRIOR APPLICATION NUMBER: 60/089532			
PRIOR FILING DATE: 1998-06-17			
PRIOR APPLICATION NUMBER: 60/089538			
PRIOR FILING DATE: 1998-06-17			
PRIOR APPLICATION NUMBER: 60/108905			
PRIOR FILING DATE: 1998-10-29			
PRIOR APPLICATION NUMBER: 60/106248			
PRIOR FILING DATE: 1998-10-28			
PRIOR APPLICATION NUMBER: 60/106178			
PRIOR FILING DATE: 1998-09-24			
PRIOR APPLICATION NUMBER: 60/101916			
PRIOR FILING DATE: 1998-09-25			
PRIOR APPLICATION NUMBER: 60/101786			
PRIOR FILING DATE: 1998-09-24			
PRIOR APPLICATION NUMBER: 60/101741			
PRIOR FILING DATE: 1998-09-23			
PRIOR APPLICATION NUMBER: 60/101738			
PRIOR FILING DATE: 1998-09-17			
PRIOR APPLICATION NUMBER: 60/101477			
PRIOR FILING DATE: 1998-09-15			
PRIOR APPLICATION NUMBER: 60/100390			
PRIOR FILING DATE: 1998-09-15			
PRIOR APPLICATION NUMBER: 60/100627			
PRIOR FILING DATE: 1998-09-16			
PRIOR APPLICATION NUMBER: 60/100848			
PRIOR FILING DATE: 1998-09-18			
PRIOR APPLICATION NUMBER: 60/100919			
PRIOR FILING DATE: 1998-09-17			
PRIOR APPLICATION NUMBER: 60/101477			
PRIOR FILING DATE: 1998-09-23			
PRIOR APPLICATION NUMBER: 60/101738			
PRIOR FILING DATE: 1998-09-24			
PRIOR APPLICATION NUMBER: 60/101741			
PRIOR FILING DATE: 1998-09-24			
PRIOR APPLICATION NUMBER: 60/101786			
PRIOR FILING DATE: 1998-09-25			
PRIOR APPLICATION NUMBER: 60/101916			
PRIOR FILING DATE: 1998-09-24			
PRIOR APPLICATION NUMBER: 60/101922			
PRIOR FILING DATE: 1998-09-24			
PRIOR APPLICATION NUMBER: 60/106178			
PRIOR FILING DATE: 1998-10-28			
PRIOR APPLICATION NUMBER: 60/106248			
PRIOR FILING DATE: 1998-10-29			
PRIOR APPLICATION NUMBER: 60/106464			
PRIOR FILING DATE: 1998-10-30			
PRIOR APPLICATION NUMBER: 60/106905			
PRIOR FILING DATE: 1998-11-03			
PRIOR APPLICATION NUMBER: 60/108787			
PRIOR FILING DATE: 1998-11-17			
PRIOR APPLICATION NUMBER: 60/108801			

PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/125259
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 60/125775
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/126773
PRIOR FILING DATE: 1999-03-29
PRIOR APPLICATION NUMBER: 60/127887
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/130232
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131022
PRIOR FILING DATE: 1999-04-26
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131445
PRIOR FILING DATE: 1999-04-28
PRIOR APPLICATION NUMBER: 60/134287
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/140650
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/140723
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146963
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/149320
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149638
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/164418
PRIOR FILING DATE: 1999-11-09
PRIOR APPLICATION NUMBER: 60/166361
PRIOR FILING DATE: 1999-11-16
PRIOR APPLICATION NUMBER: 60/169445
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169495
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169835

Query Match 93.4%; Score 2485; DB 15; Length 2683;
Best Local Similarity 97.8%; Pred. No. 0;

	Matches	2576;	Conservative	0;	Mismatches	20;	Indels	39;	Gaps	4;
QY	28	AACTTGTCCTGCGTCGCGTGAAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGAT								87
Db	50	ACCTTGAGTGCCGTCGCGTCGCGTGAAGAAATTCAGCATGGAATGTCTCTACTATTTCCTGGAT								109
QY	88	TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTTCATGATGTGCTG								147
Db	110	TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCAAACGATTTTCATGATGTGCTG								169
QY	148	GCAATGAAGAACCCTTCTGCTTACATGAGGAGCACAATCAATTAATGGCTGCTTCTG								207
Db	170	GCAATGAAGAACCCTTCTGCTTACATGAGGAGCACAATCAATTAATGGCTGCTTCTG								229
QY	208	ATGAAAATGACTGGAATGAAAACTCTACCAGTGTGGAAGCGGGAGACATGAGTGA								267
Db	230	ATGAAAATGACTGGAATGAAAACTCTACCAGTGTGGAAGCGGGAGACATGAGTGA								289
QY	268	AAAACCTCTGGAAGGAGGCGCGTGTGACGGCGTCCCTGACCACTGACAGCCCTCG								327
Db	290	AAAACCTCTGGAAGGAGGCGCGTGTGACGGCGTCCCTGACCACTGACAGCCCTCG								349
QY	328	TGGGCTCAATATATACATTTGCGGTGAACCTGATATTCCTTAGATGCCAAAAGAGATG								387
Db	350	TGGGCTCAATATATATACATTTGCGGTGAACCTGATATTCCTTAGATGCCAAAAGAGATG								409
QY	388	CCAATGGCAACATAGTCTATGAGAGAACTGCAGAAATGAGGCTGCTTATCTGCTGATC								447
Db	410	CCAATGGCAACATAGTCTATGAGAGAACTGCAGAAATGAGGCTGCTTATCTGCTGATC								469
QY	448	CATATGTTTACAACCTGACAGCATGTGTGAGAGACAGTGAACGGGGAAAAATGGCACCGGCC								507
Db	470	CGATGTTTACAACCTGACAGCATGTGTGAGAGACAGTGAACGGGGAAAAATGGCACCGGCC								529
QY	508	AAAGCCATCATTAACGCTCTTCCCTGATGGGAAACCTTTTCTCTCACACCCCGATGAGAA								567
Db	530	AAAGCCATCATTAACGCTCTTCCCTGATGGGAAACCTTTTCTCTCACACCCCGATGAGAA								589
QY	568	GATGAATTCATCTACGCTCTTCCACACACTTGTCAGTATTTCCAGAAATGGGACGAT								627
Db	590	GATGAATTCATCTACGCTCTTCCACACACTTGTCAGTATTTCCAGAAATGGGACGAT								649
QY	628	GTTCAGTGAGATTCTGTGTAACACAGCCAATGTGACACTTGGGCTCACTCATGGAAG								687
Db	650	GTTCAGTGAGATTCTGTGTAACACAGCCAATGTGACACTTGGGCTCACTCATGGAAG								709
QY	688	TGACTGTCTACAGAAACATGAGCGGGCATATGTTCCCATCGCACAAAGTGAABAATGTGT								747
Db	710	TGACTGTCTACAGAAACATGAGCGGGCATATGTTCCCATCGCACAAAGTGAABAATGTGT								769
QY	748	ACGTGTTAACAGATCAGATTCTGTGTTTGTAATACTTCCAGAAAGCATCGAAAT								807
Db	770	ACGTGTTAACAGATCAGATTCTGTGTTTGTAATACTTCCAGAAAGCATCGAAAT								829
QY	808	CATCGAGCAAACTTCC-CAAGATCTCCCATTAATGTTGATGTCTGATTCATGATC								866
Db	830	CATCGAGCAAACTTCCCTCAAGATCTCCCATTAATGTTGATGTCTGATTCATGATC								889
QY	867	CTAGCCACTTCTCAATATATTTCTACCATTAACATAAGTGAGCTTGGGATAATACTG								926
Db	890	CTAGCCACTTCTCAATATATTTCTACCATTAACATAAGTGAGCTTGGGATAATACTG								949
QY	927	GCCTGTTGTTTCCACCAATCATACTGTGAATCACAAGTATGTGCTCAATGAACTTCA								986
Db	950	GCCTGTTGTTTCCACCAATCATACTGTGAATCACAAGTATGTGCTCAATGAACTTCA								1009
QY	987	GCCTTAACCTCACTGTGAAGCTGCAGACACAGGACCTTGTCCGCCACGCCACCAAC								1046
Db	1010	GCCTTAACCTCACTGTGAAGCTGCAGACACAGGACCTTGTCCGCCACGCCACCAAC								1069
QY	1047	CCAGACCTTCAAAACCCACCCCTTCTTT-----								1074
Db	1070	CCAGACCTTCAAAACCCACCCCTTCTTTAGCAACTACTCTAAATCTTATGATTCAAACA								1129

QY 1075 ----AGGACCTGCTGGTGAACAACCCCTGGAGCTGAGTAGGATTCCCTGATGAAAACTGCC 1130
 Db 1130 CCCAGAGACCTGACTGGTGAACAACCCCTGGAGCTGAGTAGGATTCCCTGATGAAAACTGCC 1189
 QY 1131 AGATTAAACAGATATGGCCACTTTCAAGCCCATCAACAATTGTAGAGGAATCTTAGAGG 1190
 Db 1190 AGATTAAACAGATATGGCCACTTTCAAGCCCATCAACAATTGTAGAGGAATCTTAGAGG 1249
 QY 1191 TTACATCATCCAGATGACAGACGCTCTGATGCGGTGCTGATGGCCCTGAAAAAGCTCCCTAA 1250
 Db 1250 TTACATCATCCAGATGACAGACGCTCTGATGCGGTGCTGATGGCCCTGAAAAAGCTCCCTAA 1309
 QY 1251 TAGACTTTGTGTGACCTGCGCAAGGAGCATTTCCACGAGGTCTGTACCATCATTTCTG 1310
 Db 1310 TAGACTTTGTGTGACCTGCGCAAGGAGCATTTCCACGAGGTCTGTACCATCATTTCTG 1369
 QY 1311 ACCCCACCTGCGAGATCACCCAGAACAGTCTGCAGCCCTGTGATGTGATGAGATGT 1370
 Db 1370 ACCCCACCTGCGAGATCACCCAGAACAGTCTGCAGCCCTGTGATGTGATGAGATGT 1429
 QY 1371 GTCTGTGACTGTGAGACGAACCTTGATGGGTCTGGACGTACTGTGTGAACCTCACCC 1430
 Db 1430 GTCTGTGACTGTGAGACGAACCTTGATGGGTCTGGACGTACTGTGTGAACCTCACCC 1489
 QY 1431 TGGGGGATGACACAAGCCTGGCTCTCAACGACACCCCTGATTTCTGTCTGACAGAGACC 1490
 Db 1490 TGGGGGATGACACAAGCCTGGCTCTCAACGACACCCCTGATTTCTGTCTGACAGAGACC 1549
 QY 1491 CAGCCTGCGCTTTAAGATGGCAAAAGTCCCTGATCTCCGTGGCTGCTTGCCCATAT 1550
 Db 1550 CAGCCTGCGCTTTAAGATGGCAAAAGTCCCTGATCTCCGTGGCTGCTTGCCCATAT 1609
 QY 1551 TTGTCACTGTGATCTCCCTCTTGCTGTACAACAAAAACACAAGGAATACAAACCCCATAGAAA 1610
 Db 1610 TTGTCACTGTGATCTCCCTCTTGCTGTACAACAAAAACACAAGGAATACAAACCCCATAGAAA 1669
 QY 1611 AATGCTCTGGGAATGTGTGACGAACAAAGGCTGAGTGTCTTCTCAACCGTGCAAAAG 1670
 Db 1670 AATGCTCTGGGAATGTGTGACGAACAAAGGCTGAGTGTCTTCTCAACCGTGCAAAAG 1729
 QY 1671 CCGTGTCTTCTCCCGGAACAGAAAGGATCCGCTACTCAAAAAACCAAGATTAAAG 1730
 Db 1730 CCGTGTCTTCTCCCGGAACAGAAAGGATCCGCTACTCAAAAAACCAAGATTAAAG 1789
 QY 1731 GAGTTTCTTAATTTGACCTGTGTTCTGAAGCTCACTTTTCAAGTCCATTGATGTGAGA 1790
 Db 1790 GAGTTTCTTAATTTGACCTGTGTTCTGAAGCTCACTTTTCAAGTCCATTGATGTGAGA 1849
 QY 1791 TGTGCTGAGTGGCTATTAACTTTTCTTAAGATTATGTTAAATAGATATTGTGG 1850
 Db 1850 TGTGCTGAGTGGCTATTAACTTTTCTTAAGATTATGTTAAATAGATATTGTGG 1909
 QY 1851 TTGGGGAAATTGAATTTTATAGTTAAATGTCATTTTAAAGATGGAGAGGAGATA 1910
 Db 1910 TTGGGGAAATTGAATTTTATAGTTAAATGTCATTTTAAAGATGGAGAGGAGATA 1969
 QY 1911 TACTGACGAGCTTCAAGCATGTTGTAAGTGAATAAAGCAACTTAGCAAGGCTTCTT 1970
 Db 1970 TACTGACGAGCTTCAAGCATGTTGTAAGTGAATAAAGCAACTTAGCAAGGCTTCTT 2029
 QY 1971 TTCAATTATTTTATGTTCACTTATAAAGTCTTAGTAACTAGTAGATAGAAACACTG 2030
 Db 2030 TTCAATTATTTTATGTTCACTTATAAAGTCTTAGTAACTAGTAGATAGAAACACTG 2089
 QY 2031 TGTCCCGAGAGTAAGAGACAAGCTACTATTGATTAAGCCTTAACCCAGGTTAACTGCAA 2090
 Db 2090 TGTCCCGAGAGTAAGAGACAAGCTACTATTGATTAAGCCTTAACCCAGGTTAACTGCAA 2149
 QY 2091 GAAGAGCGGGATCTTCAAGCTTTCATGTAAGTGTATGCATTAAGCCAAATGTAGTCCA 2150
 Db 2150 GAAGAGCGGGATCTTCAAGCTTTCATGTAAGTGTATGCATTAAGCCAAATGTAGTCCA 2209

QY 2151 GTTCTAAGATCATGTGTTCCAGCTAAGCTAAGTAACTCCACTTCAATATACACACTCATGAACCTCC 2210
 Db 2210 GTTCTAAGATCATGTGTTCCAGCTAAGCTAAGTAACTCCACTTCAATATACACACTCATGAACCTCC 2269
 QY 2211 TGATGGAACAATAACAGGCCCCAAGCCTGTGTATGATGATGTGCACACTTGTAGACTCAGAA 2270
 Db 2270 TGATGGAACAATAACAGGCCCCAAGCCTGTGTATGATGATGTGCACACTTGTAGACTCAGAA 2329
 QY 2271 AAAATACTACTCTCATTAATGGGTGGAGTATTTTGGTGACAACTTAATTGCTTGCTG 2330
 Db 2330 AAAATACTACTCTCATTAATGGGTGGAGTATTTTGGTGACAACTTAATTGCTTGCTG 2389
 QY 2331 AGTGAAGGATGATATTATATATATTTATTTATTTCCATGAGACATTTAGTGTGCTTTT 2390
 Db 2390 AGTGAAGGATGATATTATATATTTATTTATTTCCATGAGACATTTAGTGTGCTTTT 2449
 QY 2391 TATACGAGCATGATGCTGAGTGAACCTCTTGTATATTTCCAAATTTTGTATAGTCG 2450
 Db 2450 TATACGAGCATGATGCTGAGTGAACCTCTTGTATATTTCCAAATTTTGTATAGTCG 2509
 QY 2451 CTGCACATATTTGAATCAAAATATTAAGACTTTCCAAAAATTTGCTCCCTGTTTCA 2510
 Db 2510 CTGCACATATTTGAATC-ATATATTAAGACTTTCCAAAGATGAGGTCCCTGTTTCA 2568
 QY 2511 TGGCACTTGATCAGTAAGGATTTCCCTCTGTGTTGGAACCTAAACCATTTACTATATGT 2570
 Db 2569 TGGCACTTGATCAGTAAGGATTTCACTCTGTTGTAACTAAACCATCTACTATATGT 2628
 QY 2571 TAGACAAGACATTTTCTTTCTCTCTCTGAAAAA-AAAAAGAGGGAAGAGACA 2624
 Db 2629 TAGACATGACATTTCTTTCTCTCTCTCTGAAAAAATAAGTGGGAAGAGACA 2683

RESULT 15

US-10-230-338-41
 ; Sequence 41, Application US/10230338
 ; Publication No. US20030044934A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Gerritsen, Mary
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stephan, Jean-Philippe F.
 ; APPLICANT: Watanabe, Colin L.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3530P1C92
 ; CURRENT APPLICATION NUMBER: US/10/230,338
 ; CURRENT FILING DATE: 2002-08-28
 ; PRIOR APPLICATION NUMBER: 10/119,480
 ; PRIOR FILING DATE: 2002-04-09
 ; PRIOR APPLICATION NUMBER: 60/059113
 ; PRIOR FILING DATE: 1997-09-17
 ; PRIOR APPLICATION NUMBER: 60/062287
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/063549
 ; PRIOR FILING DATE: 1997-10-28
 ; PRIOR APPLICATION NUMBER: 60/064103
 ; PRIOR FILING DATE: 1997-10-31
 ; PRIOR APPLICATION NUMBER: 60/069873
 ; PRIOR FILING DATE: 1997-12-17
 ; PRIOR APPLICATION NUMBER: 60/078910
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/079294
 ; PRIOR FILING DATE: 1998-03-25
 ; PRIOR APPLICATION NUMBER: 60/079656
 ; PRIOR FILING DATE: 1998-03-26
 ; PRIOR APPLICATION NUMBER: 60/079728


```

; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PAML.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 41
; LENGTH: 2683
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-230-338-41

```

Query Match	93.4%	Score 2485;	DB 15;	length 2683;
Best Local Similarity	97.8%;	Pred. No. 0;		
Matches 2576;	Conservative	0;	Mismatches 20;	Indels 39;
			Gaps	4;

QY	28	AACTTGGTGCCCTGCGTCCGTGAGAAATTGACGATGGAATGCTCTACTAATTTCTCTGGGAT	87
Db	50	ACCTTGAGTGCCTGCGTCCGTGAGAAATTGACGATGGAATGCTCTAATTTCTCTGGGAT	109
QY	88	TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCCAAGATTTCATGATGTGCTGG	147
Db	110	TTCTGCTCCTGGCTGCAAGATTGCCACTTGATGCCGCCCAAGATTTCATGATGTGCTGG	169
QY	148	GCAATGAAAGACCTTCTGCTTACATGAGGAGCACAATCAATTAAATGCTGCTCTTCTG	207
Db	170	GCAATGAAAGACCTTCTGCTTACATGAGGAGCACAATCAATTAAATGCTGCTCTTCTG	229
QY	208	ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGA	267
Db	230	ATGAAATGACTGGAATGAAAACTCTACCCAGTGTGGAAGCGGGGAGACATGAGGTGA	289
QY	268	AAAACTCCTGGAAGGAGGCGGTGTGCAAGCGGTCTGACCACTGACTCACCAGCCCTCG	327
Db	290	AAAACTCCTGGAAGGAGGCGGTGTGCAAGCGGTCTGACCACTGACTCACCAGCCCTCG	349
QY	328	TGGGCTCAAAATATTAACATTTGCGGTGAACCTGATATTCCCTAGATGCCAAAAAGAAATG	387
Db	350	TGGGCTCAAAATATTAACATTTGCGGTGAACCTGATATTCCCTAGATGCCAAAAAGAAATG	409
QY	388	CCAATGGCAACATAGTCTATGAGAGAAGAACTGCAGAAATGAGCGCTGTTTATCTGCTGATC	447
Db	410	CCAATGGCAACATAGTCTATGAGAGAAGAACTGCAGAAATGAGCGCTGTTTATCTGCTGATC	469
QY	448	CATATGTTTACAACCTGACACAGCATGGTCAGAGACAGTAGACGGGGAAAAATGGCACCGGCC	507
Db	470	CGATGTTTACAACCTGACACAGCATGGTCAGAGACAGTAGACGGGGAAAAATGGCACCGGCC	529
QY	508	AAAGCCATCATTAACGCTCTTCCCTGATGGGAAAACTTTTCTCTCACCAACCCCGATGGAGAA	567
Db	530	AAAGCCATCATTAACGCTCTTCCCTGATGGGAAAACTTTTCTCTCACCAACCCCGATGGAGAA	589
QY	568	GATGGAATTTCACTCTACGCTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT	627
Db	590	GATGGAATTTCACTCTACGCTCTTCCACACACTTGGTCAGTATTTCCAGAAATTTGGGACGAT	649
QY	628	GTTCAGTGAGAGTTTCTGTGTAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG	687
Db	650	GTTCAGTGAGAGTTTCTGTGTAACACAGCCCAATGTGACACTTGGGCTCAACTCATGGAAG	709
QY	688	TGACTGTCTACAGAAGACATGGACGGGCATATGTGCCATCGCACAAAGTGAAGATGTGT	747
Db	710	TGACTGTCTACAGAAGACATGGACGGGCATATGTGCCATCGCACAAAGTGAAGATGTGT	769
QY	748	ACGTGTCTACAGATCAGATTCTGTGTTTGTGACTATGTTCCAGAAGAACGATCGAAATT	807
Db	770	ACGTGTCTACAGATCAGATTCTGTGTTTGTGACTATGTTCCAGAAGAACGATCGAAATT	829
QY	808	CATCCGACGAAAACTTCC-CAAAGATCTCCCAATTATGTTTGATGTCTGTGATTCATGATC	866
Db	830	CATCCGACGAAAACTTCTCTCAAAAGATCTCCCAATTATGTTTGATGTCTGTGATTCATGATC	889
QY	867	CTAGCCACTTCTCAATTATTCTACCATTTAACAAGTGAAGCTTCGGGGATATACTG	926
Db	890	CTAGCCACTTCTCAATTATTCTACCATTTAACAAGTGAAGCTTCGGGGATATACTG	949

QY	927	GCCTGTTTGTTC	CAACCAATCATACTGTGAATCA	CACCGTATGTCTCAATGGAACCTTCA	986
Db	950	GCCTGTTTGTTC	CAACCAATCATACTGTGAATCA	CACCGTATGTCTCAATGGAACCTTCA	1009
QY	987	GCCTTAACCTCA	CTGTGAAAGCTGCAGCA	CCAGACCTTGTCCGCCACCGCACCAACCAC	1046
Db	1010	GCCTTAACCTCA	CTGTGAAAGCTGCAGCA	CCAGACCTTGTCCGCCACCGCACCAACCAC	1069
QY	1047	CCAGACCTTCA	AAACCCACCCTTCTT	-----	1074
Db	1070	CCAGACCTTCA	AAACCCACCCTTCTT	AGCAACTACTCTTAATCTTAATGATTC	1129
QY	1075	----	AGACCTGCTGTGACA	ACCCCTCGAGCTGAGTAGGATTCCTGATG	1130
Db	1130	CCCCAGGACCT	CTGTGGTGAACA	CCCCCTGAGCTGAGTAGGATTCCTGATG	1189
QY	1131	AGATTAA	CAGATATGGCCACTTTCA	AGCCCATCAATTTGAGAGGAATCTTAGAGG	1190
Db	1190	AGATTAA	CAGATATGGCCACTTTCA	AGCCCATCAATTTGAGAGGAATCTTAGAGG	1249
QY	1191	TTAATCAT	CCAGATGACAGACGTCTGATG	CCCGGTCCATGGCTGAAAGCTCCCTAA	1250
Db	1250	TTAATCAT	CCAGATGACAGACGTCTGATG	CCCGGTCCATGGCTGAAAGCTCCCTAA	1309
QY	1251	TAGACTT	GTGCTGACCTGCCA	AGGAGCATTCACGGAAGTCTGATCATCATTTCTG	1310
Db	1310	TAGACTT	GTGCTGACCTGCCA	AGGAGCATTCACGGAAGTCTGATCATCATTTCTG	1369
QY	1311	ACCCCA	CCCTGCGAGATCACCCAGAACACAGTCTG	CAACCCCTGTGGATGTGATGATGT	1370
Db	1370	ACCCCA	CCCTGCGAGATCACCCAGAACACAGTCTG	CAACCCCTGTGGATGTGATGATGT	1429
QY	1371	GTCTGCT	ACTGTGAGACGAACCTTCA	TGGGTCTGGACGTACTGTGTAACCTCA	1430
Db	1430	GTCTGCT	ACTGTGAGACGAACCTTCA	TGGGTCTGGACGTACTGTGTAACCTCA	1489
QY	1431	TGGGGG	ATGACACACAGCCTGGCTCTCA	CGACACCTTGATTTCTGTTCTTGACAGAGACC	1490
Db	1490	TGGGGG	ATGACACACAGCCTGGCTCTCA	CGACACCTTGATTTCTGTTCTTGACAGAGACC	1549
QY	1491	CAGCCT	CGCCTTAAAGATGGCAAA	CAGTGCCCTGATCTCCGTTGGCTGCTGGCCATAT	1550
Db	1550	CAGCCT	CGCCTTAAAGATGGCAAA	CAGTGCCCTGATCTCCGTTGGCTGCTGGCCATAT	1609
QY	1551	TTGTCA	CTGTGATCTCCCTCTTGGT	GTACAAAAACAAGAAATACAA	1610
Db	1610	TTGTCA	CTGTGATCTCCCTCTTGGT	GTACAAAAACAAGAAATACAA	1669
QY	1611	ATA	GTCTGGGAATGTGTGACAA	GAAGGCTGAGTCTTTCTCA	1670
Db	1670	ATA	GTCTGGGAATGTGTGACAA	GAAGGCTGAGTCTTTCTCA	1729
QY	1671	CCGT	GTCTCCCGGAAACAGAA	AAAGGATCCGCTACTCA	1730
Db	1730	CCGT	GTCTCCCGGAAACAGAA	AAAGGATCCGCTACTCA	1789
QY	1731	GAGTT	CTTAATTTGACCTGTTT	CTGAAGTCACTTTCA	1790
Db	1790	GAGTT	CTTAATTTGACCTGTTT	CTGAAGTCACTTTCA	1849
QY	1791	TGTGCT	GAGTGGCTATTAACCTTTT	TCTCAAGATTAATGTTAATAGATATG	1850
Db	1850	TGTGCT	GAGTGGCTATTAACCTTTT	TCTCAAGATTAATGTTAATAGATATG	1909
QY	1851	TTTGGG	AAATGAATTTTAAAGTTA	AGTTAATGTTAATAGATATG	1910
Db	1910	TTTGGG	AAATGAATTTTAAAGTTA	AGTTAATGTTAATAGATATG	1969
QY	1911	TACTG	CAGCAGCTTCAGCCATGTT	GTGAACCTGATTAAGCACTTAGCAAGGCTTCT	1970
Db	1970	TACTG	CAGCAGCTTCAGCCATGTT	GTGAACCTGATTAAGCACTTAGCAAGGCTTCT	2029
QY	1971	TTCA	TATTTTTTATGTTTCA	CTTAATAAGTCTTAGTA	2030

Db	2030	TTCAATATTTTATGTTTCACTTATAAGTCTTAGGTACTAGTAGATAGAACA	2089
QY	2031	TGTCCTGAGTAAGAGAGAGAGTACTATTTAGTAGAGCCCTAACCGAGTTACTGCA	2090
Db	2090	TGTCCCGAGAGTAAGAGAGAGAGTACTATTTAGTAGAGCCCTAACCGAGTTACTGCA	2149
QY	2091	GAAGAGCGGATACCTTTCAGCTTTCAGTAACTGTATGATTAAGCCCAATGTAGTCCA	2150
Db	2150	GAAGAGCGGATACCTTTCAGCTTTCAGTAACTGTATGATTAAGCCCAATGTAGTCCA	2209
QY	2151	GTTTCTAAGATCATGTTCAGAGCTTAATCCCACTTCAATACAGACTCATGAACCTCC	2210
Db	2210	GTTTCTAAGATCATGTTCAGAGCTTAATCCCACTTCAATACAGACTCATGAACCTCC	2269
QY	2211	TGATGGAACAATAACAGGCCCCAGGCTGTGTATGATGTGACACACTGTAGACTCAGAA	2270
Db	2270	TGATGGAACAATAACAGGCCCCAGGCTGTGTATGATGTGACACACTGTAGACTCAGAA	2329
QY	2271	AAAATACTACTCTCATAAATGGTGGAGTATTTTGGTGACAACTACTTGTGCTGCTG	2330
Db	2330	AAAATACTACTCTCATAAATGGTGGAGTATTTTGGTGACAACTACTTGTGCTGCTG	2389
QY	2331	AGTGAAGGAATGATATTCATATATTCATTTATTCATGACATTTAGTAGTCTTTTA	2390
Db	2390	AGTGAAGGAATGATATTCATATATTCATTTATTCATGACATTTAGTAGTCTTTTA	2449
QY	2391	TATACCAGGCATGATGCTGAGTGACACTCTGTGTATATTTCCAAATTTTGTATAGTGG	2450
Db	2450	TATACCAGGCATGATGCTGAGTGACACTCTGTGTATATTTCCAAATTTTGTATAGTGG	2509
QY	2451	CTGCACATATTTGAAATCAAAATTAAGACTTTCCAAAAATTTGTCCTGTTTTC	2510
Db	2510	CTGCACATATTTGAAATCAAAATTAAGACTTTCCAAAAATTTGTCCTGTTTTC	2568
QY	2511	TGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTGAAGCTAAACCATTACTATATGT	2570
Db	2569	TGGCAACTTGATCAGTAAGGATTTCCCTCTGTTTGAAGCTAAACCATTACTATATGT	2628
QY	2571	TAGACAAGACATTTTCTCTCTGCTGAAAAA-AAATGAGGGAAGAGACA	2624
Db	2629	TAGACATGACATTTCTCTCTCTGCTGAAAAAATAAGTGTGGAAGAGACA	2683

Search completed: September 10, 2004, 15:32:23
 Job time : 3924 secs